

STIC-Biotech/ChemLib

169038

Mej

From: Swope, Sheridan
Sent: Wednesday, October 19, 2005 5:09 PM
To: STIC-Biotech/ChemLib
Subject: 10/698,228

For 10/698,228, pls do the following:

Align SID 1 with SID 2
Align SID 1 with SID 12

Search and Interference Search:

SID 1 against the NT and AA data bases
SID 2 against the NT and AA data bases
SID 12 against the NT and AA data bases

Thanks!!

Sheridan Swope, Ph.D.
Patent Examiner, AU 1656
Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E03C70 Remsen Bld (Mailbox)

Toby Port

RECEIVED
OCT 20 2005
STIC/CHEN, LIA
(STIC)

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

STIC-Biotech/ChemLib

166062

M2

From: Swope, Sheridan
Sent: Saturday, September 17, 2005 3:04 PM
To: STIC-Biotech/ChemLib
Subject: 10/698,228

For 10/698,228,

Pls search and interference search:
SID 12 against the NT data bases.

Pls, also align SID 2 and SID 12

Sheridan Swope, Ph.D.
Patent Examiner, AU 1656
Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E03C70 Remsen Bld (Mailbox)

12 na 1053

2 na 1053

Bret Sharts

M2

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

170611
Shears, Beverly

From: Swope, Sheridan
Sent: Thursday, November 03, 2005 4:17 PM
To: Shears, Beverly
Subject: FW: 10/698,228

Importance: High

Beverly,

The results I got back indicate that the search was done for sid 2, not sid 12.
Would you do the search for sid 12 asap?

thanks,

Sheridan

-----Original Message-----

From: Swope, Sheridan
Sent: Saturday, September 17, 2005 3:04 PM
To: STIC-Biotech/ChemLib
Subject: 10/698,228

Bey Shears

For 10/698,228,

Pls search and interference search:
SID 12 against the NT data bases.

Pls, also align SID 2 and SID 12

19280

Sheridan Swope, Ph.D.
Patent Examiner, AU 1656
Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E03C70 Remsen Bld (Mailbox)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2005, 06:51:46 ; Search time 672 Seconds
(without alignments)
9276.023 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1053

Sequence: 1 atggcgaccgcgaaagcc.....tcaccttgatgcagatttt 1053

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1053	100.0	1053	4	Aaf32438 Human rib
2	1053	100.0	2596	3	Aa12411 cDNA enco
3	1053	100.0	4955	4	Aaf32439 Human rib
4	1053	100.0	4955	13	ADR24210 Breast ca
5	1051.4	99.8	1053	4	Aaf32447 Human rib
6	1051.4	99.8	1081	4	Aaf32440 Human rib
7	1049.8	99.7	1601	4	Aa114924 Human cDN
8	650.6	61.8	706	13	ADQ57092 Novel can
9	578.4	54.9	1989	3	Aac78111 Human can
10	578.4	54.9	2215	12	ADK70302 Respirato
11	578.4	54.9	2216	10	ADJ56536 Murine cD
12	578.4	54.9	2482	4	Aa544917 Human con
13	578.4	54.9	2500	6	ABL65414 Lung can
14	578.4	54.9	2500	6	ABL65517 Lung can
15	578.4	54.9	2500	6	ABL65859 Lung can
16	578.4	54.9	2500	8	ABX10335 DNA enco
17	578.4	54.9	2500	11	ADJ32044 Human cDN
18	578.4	54.9	2500	12	ADN04443 Antipsoi
19	578.4	54.9	2500	12	ADQ09273 Human RRM
20	578.4	54.9	2500	13	ACN37637 Tumour-as

21	578.4	54.9	3294	4	AAH73225	Aah73225 Human cer
22	578.4	54.9	3393	12	ADN03788	Adn03788 Antipsoi
23	578.4	54.9	3393	12	ADO19225	Ado19225 Human PRO
24	569	54.0	2641	4	AAS44745	Aas44745 Human ful
25	568.8	54.0	1170	2	AAV05641	Aav05641 Human rib
26	561	53.3	1328	12	ADO57308	Ado57308 DNA enco
27	547.8	52.0	978	4	AAS41006	Aas41006 cDNA enco
28	545.4	51.8	977	6	ABL90228	Ab190228 Human pol
29	538.8	50.2	608	4	AAH07707	Aah07707 Human cDN
30	514	48.8	1371	5	AAH79474	Aah79474 DNA enco
31	463	44.0	1289	4	ABL14627	Ab114627 Drosophil
32	442	42.0	1146	6	AAS62387	Aas62387 cDNA sequ
33	441	41.9	481	2	AAS51874	Aas51874 Human sec
34	432.2	41.0	1218	8	ABT20895	Abt20895 Aspergill
35	429.8	40.8	14176	2	AAH84564	Aah84564 Swinpo
C	429.8	40.8	14176	4	AAF84949	Aaf84949 Reverse c
36	429.8	40.8	14176	4	AAF84948	Aaf84948 Nucleotid
C	420.4	39.9	3945	4	ABL14626	Ab114626 Drosophil
38	403.2	38.3	963	13	ADS58378	AdS58378 Bacteri
39	397	37.7	1146	8	ABT19075	Abt19075 Aspergill
40	392.6	37.3	1173	13	ADT47814	Adt47814 Bacteri
41	392.6	37.3	1242	6	ABZ32250	Abz32250 Candida a
42	392	37.2	1292	13	ADS49379	AdS49379 Bacteri
43	386.6	36.7	1292	13	ACC61238	Acc61238 Gene sequ
44	380.4	36.1	1200	10	ADK63241	Adk63241 Disease t
45	380.4	36.1	1200	10	ADK63241	Adk63241 Disease t

ALIGNMENTS

RESULT 1

AAF32438

ID AAF32438 standard; cDNA; 1053 BP.

AC AAF32438;

XX 18-APR-2001 (first entry)

DE Human ribonucleotide reductase TP53R2H nucleotide sequence SEQ ID NO:2.

XX Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.

OS Homo sapiens.

XX WO200100799-A1.

PD 04-JAN-2001.

XX 27-JUN-2000; 2000WO-JP004189.

XX 28-JUN-1999; 99JP-00181131.

PR 06-JUL-1999; 99JP-00192391.

PR 21-JAN-2000; 2000JP-00017770.

XX (TAKE) TAKEDA CHEM IND LTD.

PA (NAKA/) NAKAMURA Y.

XX Nakamura Y, Arakawa H, Tanaka H;

XX WPI; 2001-112446/12.

DR P-PSDB; AAB69050.

XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for

PT diagnosis, treatment and prevention of cancer.

XX Claim 5; Fig 1-3; 102pp; Japanese.

CC The present sequence encodes a human ribonucleotide reductase designated

CC TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of the DNA

CC repair mechanism and its activity is induced by p53. It can be used for

CC the treatment, prevention and diagnosis of a wide range of cancers

XX Sequence 1053 BP; 314 A; 192 C; 246 G; 301 T; 0 U; 0 Other;

		Query Match	100.0%; Score 1053; DB 4; Length 1053;	
		Best Local Similarity	100.0%; Pred. NO. 1e-290;	
		Mismatches 1053; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	ATGGCGGACCCGGAAGCGCGGAGCGGGCTGGATCAGATCAGAGATCATCTTCA	60	
DB	1	ATGGCGGACCCGGAAGCGCGGAGCGGGCTGGATCAGATCAGAGATCATCTTCA	60	
QY	61	GACACCAACGAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAAGAGTTCCTGC	120	
DB	61	GACACCAACGAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAAGAGTTCCTGC	120	
QY	121	CGGTTTGTGCATCTTTCCATCCAGTACCCTGATATTTGGAAATGATATAACAGGCACAG	180	
DB	121	CGGTTTGTGCATCTTTCCATCCAGTACCCTGATATTTGGAAATGATATAACAGGCACAG	180	
QY	181	GCTTCCTTCTGACAGCAGAGAGTGCAGTATCAAGAGTCCCTCCTCAGTGAACAAG	240	
DB	181	GCTTCCTTCTGACAGCAGAGAGTGCAGTATCAAGAGTCCCTCCTCAGTGAACAAG	240	
QY	241	CTTAAAGCAGATGAGAAGTACTTCTCTCTCATCTCTCATCTTTAGCCCTTTTTCAGCCAGTGAT	300	
DB	241	CTTAAAGCAGATGAGAAGTACTTCTCTCTCATCTCTCATCTTTAGCCCTTTTTCAGCCAGTGAT	300	
QY	301	GGAAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGAGGTTCCAGAGGCT	360	
DB	301	GGAAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGAGGTTCCAGAGGCT	360	
QY	361	CGCTGTTTCTATGGCTTCAAAATTCATCGAGAAATGTTTCACTCAGAGATGATACAGTTTG	420	
DB	361	CGCTGTTTCTATGGCTTCAAAATTCATCGAGAAATGTTTCACTCAGAGATGATACAGTTTG	420	
QY	421	CTGATAGACACTTACATCAGATGATCCAGAAAGGGAATTTTATTAATGCAATTTGAA	480	
DB	421	CTGATAGACACTTACATCAGATGATCCAGAAAGGGAATTTTATTAATGCAATTTGAA	480	
QY	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGATAGCAGATAGAAAA	540	
DB	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGATAGCAGATAGAAAA	540	
QY	541	TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAAGAGAGTTCCTCTCAGGA	600	
DB	541	TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAAGAGAGTTCCTCTCAGGA	600	
QY	601	TCTTTTGTCTATATTTCTGCTTAAGAGAGAGTCTTATGCGAGACTCACTTTTTC	660	
DB	601	TCTTTTGTCTATATTTCTGCTTAAGAGAGAGTCTTATGCGAGACTCACTTTTTC	660	
QY	661	AATGAACTCATCAGCAGATGAAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA	720	
DB	661	AATGAACTCATCAGCAGATGAAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA	720	
QY	721	TACTTAGTAATAAGCTTTCAGAAAGAGGTGAGGAGATCATTTGATGCTGCAAA	780	
DB	721	TACTTAGTAATAAGCTTTCAGAAAGAGGTGAGGAGATCATTTGATGCTGCAAA	780	
QY	781	ATTGACGAGAGTTTAAACAGAGCTTGCAGTTGGCTCATTTGGAATGAATTCATT	840	
DB	781	ATTGACGAGAGTTTAAACAGAGCTTGCAGTTGGCTCATTTGGAATGAATTCATT	840	
QY	841	TTGATGAAACAGTACATGATTTGAGCTGACAGATTAATCTTGGAACTTGGATTTCTCA	900	
DB	841	TTGATGAAACAGTACATGATTTGAGCTGACAGATTAATCTTGGAACTTGGATTTCTCA	900	
QY	901	AAGGTTTTTTCAGGACAGAAATCCTTTTGTATTTTATGAAACATTTCTTTAGAGGAAA	960	
DB	901	AAGGTTTTTTCAGGACAGAAATCCTTTTGTATTTTATGAAACATTTCTTTAGAGGAAA	960	
QY	961	ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGGATATGCGAGAAACC	1020	
DB	961	ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGGATATGCGAGAAACC	1020	

QY 1021 ACAGATAACGCTCTTACCTTGGATCGAGATTTT 1053
DB 1021 ACAGATAACGCTCTTACCTTGGATCGAGATTTT 1053

RESULT 2
AA12411
ID AAA12411 standard; cDNA; 2596 BP.
XX
AC AAA12411;
XX
XX 25-JUL-2000 (first entry)
XX
DE cDNA encoding a human RNA-associated protein.

XX Human; RNA-associated protein; cell proliferation; cancer; inflammation;
XX immune response; reproductive disorder; actinic keratosis;
KW atherosclerosis; arteriosclerosis; bursitis; cirrhosis; hepatitis;
KW mixed connective tissue disease; myelofibrosis; primary thrombocythemia;
KW paroxysmal nocturnal hemoglobinuria; polycythemia vera; psoriasis;
KW trauma; ss.
XX
XX Homo sapiens.

XX Key Location/Qualifiers
FH CDS 56..1111
FT FT /*tag= a
FT FT /product= "RNA-associated protein"
XX
XX WO200015799-A2.

XX 23-MAR-2000.

XX 17-SEP-1999; 99WO-US021688.
XX
XX 17-SEP-1998; 98US-00156039.
PR 22-SEP-1998; 98US-00158720.
PR 04-NOV-1998; 98US-00186815.
PR 08-APR-1999; 99US-0128660P.

XX (INCY-) INCYTE PHARM INC.
XX
XX Tang YT, Corley NC, Guegler KJ, Gorgone GA, Patterson C;
PI Hillman JL, Baughn MR, Lal P, Azimzai Y, Yue H, Yang J;
XX
XX WPI: 2000-271437/23.
DR P-PSDB; AAY84439.

XX New polypeptides and polynucleotides, useful for preventing and treating
PT a disorder associated with increased or decreased expression of RNA
PT associated proteins.
XX
XX Claim 9; Page 120-121; 131pp; English.
XX
XX The present sequence encodes a human RNA-associated protein. The
CC expression of RNA-associated proteins is closely associated with
CC reproductive tissues, nervous tissues, cell proliferation including
CC cancer, inflammation and immune responses, and so they may be used for
CC diagnosis, treatment or prevention of cell proliferative,
CC immune/inflammatory disorders, and reproductive disorders. Diseases and
CC disorders which may be treated include actinic keratosis,
CC atherosclerosis, arteriosclerosis, bursitis, cirrhosis, hepatitis, mixed
CC connective tissue disease, myelofibrosis, paroxysmal nocturnal
CC hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia
CC and cancers, and trauma
XX
XX Sequence 2596 BP; 783 A; 418 C; 510 G; 885 T; 0 U; 0 Other;

Query Match 100.0%; Score 1053; DB 3; Length 2596;
Best Local Similarity 100.0%; Pred. NO. 1.5e-290;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGGACCCGGAAGCGCGGAGCGGGCTGGATCAGATGAGATCATCTTCA 60

Db 56 ATGGGCGACCCGGAAGGCGGAGCGGCGGCTGGATCAGATGAGATCATCTTCA 115
 Qy 61 GACACCAACGAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
 Db 116 GACACCAACGAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGAGTTCTCGC 175
 Qy 121 CGGTTTGTCTATCTTCCATCCAGTACCCTTGATATTTGGAAATGTATTAACAGGCAAG 180
 Db 176 CGGTTTGTCTATCTTCCATCCAGTACCCTTGATATTTGGAAATGTATTAACAGGCAAG 235
 Qy 181 GCTTCTCTTGGACAGCAGAGAGGTGCGACTTATCAAGAGATCTCCTCACTGCAACAAG 240
 Db 236 GCTTCTCTTGGACAGCAGAGAGGTGCGACTTATCAAGAGATCTCCTCACTGCAACAAG 295
 Qy 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 300
 Db 296 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 355
 Qy 301 GGAATTTGTAATGAATAATTTGGAGCGCTTTAGTCAGAGGTGCAAGTTCCAGAGGCT 360
 Db 356 GGAATTTGTAATGAATAATTTGGAGCGCTTTAGTCAGAGGTGCAAGTTCCAGAGGCT 415
 Qy 361 CGCTGTTTCTATGGCTTCAATTTCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 420
 Db 416 CGCTGTTTCTATGGCTTCAATTTCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 475
 Qy 421 CTGATAGACACTTATCATCAGAGATCCCAAGAAAGGAATTTTATTAATGCAATTGAA 480
 Db 476 CTGATAGACACTTATCATCAGAGATCCCAAGAAAGGAATTTTATTAATGCAATTGAA 535
 Qy 481 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGGCCCTTGCATGATAGCAGATAGAAAA 540
 Db 536 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGGCCCTTGCATGATAGCAGATAGAAAA 595
 Qy 541 TCTACTTTTGGGAAAGAGTGGGCTTTGCTGTGTAGAGAGATTTTCTTCTCAGGA 600
 Db 596 TCTACTTTTGGGAAAGAGTGGGCTTTGCTGTGTAGAGAGATTTTCTTCTCAGGA 655
 Qy 601 TCTTTTGTCTATATCTGGCTTAAGAGAGAGGTCTTATGCGAGACTCACTTTTTC 660
 Db 656 TCTTTTGTCTATATCTGGCTTAAGAGAGAGGTCTTATGCGAGACTCACTTTTTC 715
 Qy 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
 Db 716 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 775
 Qy 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTGAGGAGATCATTTGTGATGCTGTCAAA 780
 Db 776 TACTTAGTAAATAAGCCTTCAGAAAGAGGTGAGGAGATCATTTGTGATGCTGTCAAA 835
 Qy 781 ATTGACGAGAGTTTAAACAGAGAGCTTGCAGAGTTGGCCTCATTTGGAATGCAATTGCAAT 840
 Db 836 ATTGACGAGAGTTTAAACAGAGAGCTTGCAGAGTTGGCCTCATTTGGAATGCAATTGCAAT 895
 Qy 841 TTGATCAAAACAGTACATGAGTTTGTAGCTGACAGATTAATCTGTGAACTTTGATTTCTCA 900
 Db 896 TTGATCAAAACAGTACATGAGTTTGTAGCTGACAGATTAATCTGTGAACTTTGATTTCTCA 955
 Qy 901 AAGGTTTTTTCAGGCAAGAAATCCTTTTGTATTTTATGGAACATTTCTTTAGAGGAAAA 960
 Db 956 AAGGTTTTTTCAGGCAAGAAATCCTTTTGTATTTATGGAACATTTCTTTAGAGGAAAA 1015
 Qy 961 ACAATTTCTTTGAGAAAACAGTTTTCAGAGTATCAGCGTTTTCAGAGTTTTCAGGCAAGAAC 1020
 Db 1016 ACAATTTCTTTGAGAAAACAGTTTTCAGAGTATCAGCGTTTTCAGAGTTTTCAGGCAAGAAC 1075
 Qy 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
 Db 1076 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1108

AAF32439
 ID AAF32439 standard; cDNA; 4955 BP.
 XX
 AC AAF32439;
 DT 18-APR-2001 (first entry)
 XX
 DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:3.
 XX
 KW Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
 XX
 OS Homo sapiens.
 XX
 FN WO2001007999-A1.
 XX
 PD 04-JAN-2001.
 XX
 PF 27-JUN-2000; 2000WO-JP004189.
 XX
 PR 28-JUN-1999; 99JP-00181131.
 PR 06-JUL-1999; 99JP-00192391.
 PR 21-JAN-2000; 2000JP-00017770.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 PA (NAKA/) NAKAMURA Y.
 PI Nakamura Y, Arakawa H, Tanaka H;
 XX
 WIPI; 2001-112446/12.
 XX
 PT Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
 PT diagnosis, treatment and prevention of cancer.
 XX
 PS Example 2; Page 87-90; 102pp; Japanese.
 XX
 CC The present invention describes a human ribonucleotide reductase
 CC designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of
 CC the DNA repair mechanism and its activity is induced by p53. It can be
 CC used for the treatment, prevention and diagnosis of a wide range of
 CC cancers. The present sequence represents a human ribonucleotide reductase
 CC related sequence which is used in an example from the present invention
 XX
 SQ Sequence 4955 BP; 1526 A; 803 C; 999 G; 1627 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1053; DB 4; Length 4955;
 Best Local Similarity 100.0%; Pred. No. 2.1e-250;
 Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ATGGGCGACCCGGAAGGCGGAGCGGCGGCTGGATCAGATGAGATCATCTTCA 60
 Db 245 ATGGGCGACCCGGAAGGCGGAGCGGCGGCTGGATCAGATGAGATCATCTTCA 304
 Qy 61 GACACCAACGAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
 Db 305 GACACCAACGAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
 Qy 121 CGGTTTGTCTATCTTCCATCCAGTACCCTTGATATTTGGAAATGTATTAACAGGCAAG 180
 Db 365 CGGTTTGTCTATCTTCCATCCAGTACCCTTGATATTTGGAAATGTATTAACAGGCAAG 424
 Qy 181 GCTTCTCTTGGACAGCAGAGAGGTGCGACTTATCAAGAGATCTCCTCACTGGAACAAG 240
 Db 425 GCTTCTCTTGGACAGCAGAGAGAGGTGCGACTTATCAAGAGATCTCCTCACTGGAACAAG 484
 Qy 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 300
 Db 485 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 544
 Qy 301 GGAATTTGTAATGAATAATTTGGAGCGCTTTAGTCAGAGGTGCAAGTTCCAGAGGCT 360
 Db 545 GGAATTTGTAATGAATAATTTGGAGCGCTTTAGTCAGAGGTGCAAGTTCCAGAGGCT 604
 Qy 361 CGCTGTTTCTATGGCTTCAATTTCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 420

Db 605 CGCTGTTCTATGGCTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 664
 QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 480
 Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 724
 QY 481 ACATGCCCTATGTTAAGAAAAAAGCAGATGTGGCCCTTGCATGGAATAGAGATAAAAA 540
 Db 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATGTGGCCCTTGCATGGAATAGAGATAAAAA 784
 QY 541 TCTACTTTTGGGAAGAGTGGGCTTGGCTGCTGTAGAGAGAGTTCCTCTCAGGA 600
 Db 785 TCTACTTTTGGGAAGAGTGGGCTTGGCTGCTGTAGAGAGAGTTCCTCTCAGGA 844
 QY 601 TCTTTTGGCTGCTATATCTGGCTTAAAGAAAGAGAGTCTTATGCGAGGACTCCTTTTTC 660
 Db 845 TCTTTTGGCTGCTATATCTGGCTTAAAGAAAGAGAGTCTTATGCGAGGACTCCTTTTTC 904
 QY 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
 Db 905 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
 QY 721 TACTTAGTAATAAGCCTTCAGAAAGAGGCTCAGGAGATCAATGTTGATGCTGTCAAA 780
 Db 965 TACTTAGTAATAAGCCTTCAGAAAGAGGCTCAGGAGATCAATGTTGATGCTGTCAAA 1024
 QY 781 ATTGACGAGAGTTTTAAACAGAAAGCCTTCCAGTTGGCTCAATGGAATGAATGCAAT 840
 Db 1025 ATTGACGAGAGTTTTAAACAGAAAGCCTTCCAGTTGGCTCAATGGAATGAATGCAAT 1084
 QY 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTTCTCA 900
 Db 1085 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTTCTCA 1144
 QY 901 AAGTTTTTTCAGGACAGAAATCCTTTGATTTATGGAACCAATTTCTTTAGAGGAAAA 960
 Db 1145 AAGTTTTTTCAGGACAGAAATCCTTTGATTTATGGAACCAATTTCTTTAGAGGAAAA 1204
 QY 961 ACAATTTTCTTTCAGAAAGAGTTTCAGAGTATCAGGTTTTCAGGTTATGCGAGAAACC 1020
 Db 1205 ACAATTTTCTTTCAGAAAGAGTTTCAGAGTATCAGGTTTTCAGGTTATGCGAGAAACC 1264
 QY 1021 ACAGATAAGCTCTTACCTTTGGATGCGAGTTT 1053
 Db 1265 ACAGATAAGCTCTTACCTTTGGATGCGAGTTT 1297

RESULT 4
 ADR24210
 ID ADR24210 standard; DNA; 4955 BP.
 XX AC ADR24210;
 XX DT 21-OCT-2004 (first entry)
 XX DE Breast cancer prognosis marker #71.
 XX KW ds; breast cancer; prognosis; gene expression; diagnosis.
 XX OS Homo sapiens.
 XX PN WO2004065545-A2.
 XX PD 05-AUG-2004.
 XX PF 15-JAN-2004; 2004WO-US001100.
 XX PR 15-JAN-2003; 2003US-00342887.
 XX PA (ROSE-) ROSETTA INPHARMATICS LLC.
 XX PA (NECA-) NETHERLANDS CANCER INST.

PI Van't Veer LJ, He Y;
 DR WPI; 2004-593473/57.
 XX Classifying a breast cancer patient according to prognosis comprises
 PT determining the similarity between the level of expression of each of
 PT five genes in a cell sample taken from patient, to control levels.
 XX Disclosure; SEQ ID NO 71; 226pp; English.
 XX The invention relates to a method of classifying a breast cancer patient
 CC according to prognosis by determining the similarity between the level of
 CC expression of each of five genes for which markers are listed in the
 CC specification, in a cell sample taken from the breast cancer patient, to
 CC control levels of expression for each respective five genes to obtain a
 CC patient similarity value. The methods are useful for classifying a breast
 CC cancer patient according to prognosis. Kits and computer program products
 CC are useful for data analysis using the diagnostic, prognostic and
 CC statistical methods of the invention. This sequence corresponds to a
 CC marker used in the method of the invention.
 XX SQ Sequence 4955 BP; 1526 A; 803 C; 999 G; 1627 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1053; DB 13; Length 4955;
 Best Local Similarity 100.0%; Pred. No. 2.1e-290;
 Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGCGGACCCGGAAGGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
 Db 245 ATGGCGGACCCGGAAGGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
 QY 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCG 120
 Db 305 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCG 364
 QY 121 CGGTTGTGTCATCTTTCCAAATCCAGTACCTGATTTTGGAAAATGTATAAACAGGCACAG 180
 Db 365 CGGTTGTGTCATCTTTCCAAATCCAGTACCTGATTTTGGAAAATGTATAAACAGGCACAG 424
 QY 181 GCTTCTCTTCGACAGCAGAGAGTGCGACTTATCAAAAGGATCTCCCTCAGTGAACAAG 240
 Db 425 GCTTCTCTTCGACAGCAGAGAGTGCGACTTATCAAAAGGATCTCCCTCAGTGAACAAG 484
 QY 241 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTTAGCCTTTTTCAGCCAGTGAT 300
 Db 485 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTTAGCCTTTTTCAGCCAGTGAT 544
 QY 301 GGAATTGTAAATGAAAAATTTGGTGAGCGCTTTTAGTCAGGAGTGCGAGTTCCAGAGGCT 360
 Db 545 GGAATTGTAAATGAAAAATTTGGTGAGCGCTTTTAGTCAGGAGTGCGAGTTCCAGAGGCT 604
 QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGATGTTCTCATCAGAGATGTACAGTTTG 420
 Db 605 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGATGTTCTCATCAGAGATGTACAGTTTG 664
 QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 480
 Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 724
 QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCTTTCGATGATGATGATGATGATGAT 540
 Db 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCTTTCGATGATGATGATGATGATGAT 784
 QY 541 TCTACTTTTGGGAAGAGTGGGCTTTCCTGCTGTAGAGAGTTCCTCTCAGGA 600
 Db 785 TCTACTTTTGGGAAGAGTGGGCTTTCCTGCTGTAGAGAGTTCCTCTCAGGA 844
 QY 601 TCTTTTGGCTGCTATATCTGCTAAAGAGAGAGTCTTATGCGAGGACTCCTTTTTC 660
 Db 845 TCTTTTGGCTGCTATATCTGCTAAAGAGAGAGTCTTATGCGAGGACTCCTTTTTC 904
 QY 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720


```
Db 1021 ACAGATAACGCTCTTCACTCTGGATGAGATTTT 1053
|||||
RESULT 6
AAF32440
ID AAF32440 standard; cDNA; 1081 BP.
XX
AC AAF32440;
XX
DT 18-APR-2001 (first entry)
XX
DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:4.
XX
KW Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
XX
OS Homo sapiens.
XX
XX WO200100799-A1.
XX
PD 04-JAN-2001.
XX
PF 27-JUN-2000; 2000WO-JP004189.
XX
PR 28-JUN-1999; 99JP-00181131.
XX
PR 06-JUL-1999; 99JP-00192391.
XX
PR 21-JAN-2000; 2000JP-00017770.
XX
XX (TAKE ) TAKEDA CHEM IND LTD.
XX
PA (NARA/) NAKAMURA Y.
XX
PI Nakamura Y, Arakawa H, Tanaka H;
XX
XX WPI; 2001-112446/12.
XX
DR Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
XX PT diagnosis, treatment and prevention of cancer.
XX PS Example 2; Page 91; 102pp; Japanese.
XX
XX The present invention describes a human ribonucleotide reductase
CC designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of
CC the DNA repair mechanism and its activity is induced by p53. It can be
CC used for the treatment, prevention and diagnosis of a wide range of
CC cancers. The present sequence represents a human ribonucleotide reductase
CC related sequence which is used in an example from the present invention
XX
XX Sequence 1081 BP; 322 A; 199 C; 252 G; 308 T; 0 U; 0 Other;
XX
Query Match 99.8%; Score 1051.4; DB 4; Length 1081;
Best Local Similarity 99.9%; Pred. No. 2.9e-290;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGGCGACCCGGAAGGCGGAGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 60
DB 20 ATGGGCGACCCGGAAGGCGGAGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 79
QY 61 GACACCAAGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGTCTTCGC 120
DB 80 GACACCAAGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGTCTTCGC 139
QY 121 CGGTTTGTGATCTTTCCATCCAGTACCTGATATTGGAAATGTATAAACAGGCACAG 180
DB 140 CGGTTTGTGATCTTTCCATCCAGTACCTGATATTGGAAATGTATAAACAGGCACAG 199
QY 181 GCTTCTCTTGGACAGCAGAGAGGTCGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
DB 200 GCTTCTCTTGGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 259
QY 241 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCAATCTAGCCCTTTTTCGAGCCAGTGAT 300
DB 260 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCAATCTAGCCCTTTTTCGAGCCAGTGAT 319
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGCT 360
DB 320 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGCT 379
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 420
DB 380 CGCTGTTTCTATGGCTTTCAAAATCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 439
QY 421 CTGATAGACACTTATCATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATGAA 480
DB 440 CTGATAGACACTTATCATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATGAA 499
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTCCGATGGATAGCAGATAGAAA 540
DB 500 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTCCGATGGATAGCAGATAGAAA 559
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTCTGCTGTAGAAAGGAGTTTCTTCTCAGGA 600
DB 560 TCTACTTTTGGGAAAGAGTGGTGGCTTCTGCTGTAGAAAGGAGTTTCTTCTCAGGA 619
QY 601 TCTTTTGTCTCTATATTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
DB 620 TCTTTTGTCTCTATATTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTCC 679
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGATGTTCCA 720
DB 680 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGATGTTCCA 739
QY 721 TACTTAGTAAATGAAGCTTTCAGAAAGAAAGGCTCAGGAGAGATCATTTGTGATGCTGCA 780
DB 740 TACTTAGTAAATGAAGCTTTCAGAAAGAAAGGCTCAGGAGAGATCATTTGTGATGCTGCA 799
QY 781 ATTGAGCAGGAGTTTAAACAGAAAGCCCTTGGCAGTTGGCCCTCATTTGGAATGAAATGCA 840
DB 800 ATTGAGCAGGAGTTTAAACAGAAAGCCCTTGGCAGTTGGCCCTCATTTGGAATGAAATGCA 859
QY 841 TTGATGAAAACAGTACATGAGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGATCTCA 900
DB 860 TTGATGAAAACAGTACATGAGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGATCTCA 919
QY 901 AAGGTTTTTTCAGGCAGAGAAAATCCCTTTTGATTTTATGGAATAACATTTCTTTAGAAGAAA 960
DB 920 AAGGTTTTTTCAGGCAGAGAAAATCCCTTTTGATTTTATGGAATAACATTTCTTTAGAAGAAA 979
QY 961 ACAAAATTTCTTTGAGAAAACGAGTTTTCAGAGTATCAGCGCTTTTTCAGAGTTATGCGAGAAACC 1020
DB 980 ACAAAATTTCTTTGAGAAAACGAGTTTTCAGAGTATCAGCGCTTTTTCAGAGTTATGCGAGAAACC 1039
QY 1021 ACAGATAACGCTCTTCACTCTGGATGAGATTTT 1053
DB 1040 ACAGATAACGCTCTTCACTCTGGATGAGATTTT 1072
XX
RESULT 7
AAH14924
ID AAH14924 standard; cDNA; 1601 BP.
XX
AC AAH14924;
XX
XX 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:12810.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-00116126.
XX
```


PR 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX (HELI-) HELIX RES INST.
 PA
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 PS Claim 8; SEQ ID NO 12810; 2537pp + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 XX
 SQ Sequence 1601 BP; 484 A; 290 C; 348 G; 479 T; 0 U; 0 Other;

QY 361 CGCTGTTTCTATGCGCTTTCAAATTTCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 420
 Db 400 CGCTGTTTCTATGCGCTTTCAAATTTCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 459
 QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATCAATGAA 480
 Db 460 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATCAATGAA 519
 QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTGGGCTTTCGATGAGATGAGATGAGAAA 540
 Db 520 ACCATGCCCTATGTTAAGAAAAGCAGATTGGGCTTTCGATGAGATGAGATGAGAAA 579
 QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTCTGCTGTAGAGAGAGTTTCTTCTCAGGA 600
 Db 580 TCTACTTTTGGGGAAGAGTGGTGGCTTCTGCTGTAGAGAGAGTTTCTTCTCAGGA 639
 QY 601 TCTTTTGTCTATATCTTGGCTAAGAGAGAGGTCTTATGCCAGGACTCATCTTTTTC 660
 Db 640 TCTTTTGTCTATATCTTGGCTAAGAGAGAGGTCTTATGCCAGGACTCATCTTTTTC 699
 QY 661 AATGAATCATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGATGTTCCAA 720
 Db 700 AATGAATCATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGATGTTCCAA 759
 QY 721 TACTTAGTAATAAGCCTTTCAGAGAAAAGGTTCAGGAGATCATTTGTTGATGCTGTCAA 780
 Db 760 TACTTAGTAATAAGCCTTTCAGAGAAAAGGTTCAGGAGATCATTTGTTGATGCTGTCAA 819
 QY 781 ATTGAGCAGGAGTTTAAACAGAGAGGCTTGCAGTTCAGTTCATTTGGAATGATTCATTT 840
 Db 820 ATTGAGCAGGAGTTTAAACAGAGAGGCTTGCAGTTCAGTTCATTTGGAATGATTCATTT 879
 QY 841 TTGATGAACAAGTACATTTGAGTTTGTAGCTGCACAGATTACTTGTGGAACCTTGGATTCTCA 900
 Db 880 TTGATGAACAAGTACATTTGAGTTTGTAGCTGCACAGATTACTTGTGGAACCTTGGATTCTCA 939
 QY 901 AAGGTTTTTTCAGGCAAGAAAATCCTTTTGAATTTTATGGAACAAATTTCTTTAGAGGAAA 960
 Db 940 AAGGTTTTTTCAGGCAAGAAAATCCTTTTGAATTTTATGGAACAAATTTCTTTAGAGGAAA 999
 QY 961 ACAATTTTCTTTGAGAAAACGAGTTTCAGAGTATCAGCGTTTTCGAGTTATGCGAGAAACC 1020
 Db 1000 ACAATTTTCTTTGAGAAAACGAGTTTCAGAGTATCAGCGTTTTCGAGTTATGCGAGAAACC 1059
 QY 1021 ACAGATACGCTTTCACCTTCGATGAGATTTT 1053
 Db 1060 ACAGATACGCTTTCACCTTCGATGAGATTTT 1092

RESULT 8
 ADQ57092
 ID ADQ57092 standard; DNA; 706 BP.
 XX
 XX ADQ57092;
 XX
 XX
 DT 21-OCT-2004 (first entry)
 XX
 DE Novel canine microarray-related DNA sequence SeqID8394.
 XX
 XX canine microarray; drug screening; toxicity assay;
 KW environmental pollutant; cellular response; gene expression profile;
 KW toxic response; liver necrosis; fatty liver disease;
 KW protein adduct formation; hepatitis; dogs; db.
 XX
 OS Canis familiaris.
 XX
 EN WO2004063324-A2.
 XX
 XX 29-JUL-2004.
 PD
 XX 05-MAY-2003; 2003WO-US013853.
 XX
 XX 03-MAY-2002; 2002US-0377240P.
 PR

XX	(GENE-) GENE LOGIC INC.	709	CTGATGTTCCATTA	768
PA	(PFIZ) PFIZER PROD INC.	361	CTGATGTTTCAATA	420
XX	Diggins JC, Porter M, Wei T;	769	GATGCTGTCAAAAT	828
PI	WPI; 2004-561890/54.	421	AATGCTGTGAAAT	480
DR		829	ATGAATTTGCATTT	888
XX	New isolated nucleic acid molecule, useful for drug screening and	481	ATGAATTTGTTGT	540
PT	toxicity assays or for assessing the impact, including toxicity, of a	889	CTTGGAATTTCTCAA	948
PT	compound, pharmaceutical agent or environmental pollutant on a cell or	541	CTTGGAATTTCTCAA	600
XX	living organism.	949	TTAGAGGAAACAAAT	1008
PS	Claim 1; SEQ ID NO 8394; 41pp; English.	601	TTAGAGGAAACAAAT	660
XX	This invention is related to a novel isolated canine nucleic acid	1009	ATGCGCAGAAACAC	1053
CC	sequences and the construction of canine microarrays containing a	661	ATGCGCAGAAACAC	705
CC	significant portion of the canine genome. The isolated canine nucleic			
CC	acid sequences of the invention may be useful for drug screening and			
CC	toxicity assays. The invention is therefore useful for assessing the			
CC	impact, including toxicity, of a compound, pharmaceutical agent or			
CC	environmental pollutant on a cell or living organism. The methods are			
CC	useful for detecting genes that are up- or down-regulated in canines in a			
CC	disease state. The sequences are useful as diagnostic agents or markers			
CC	to detect a cellular response in a sample individually or as part of a			
CC	gene expression profile. It is also useful as a target for agents that			
CC	modulate gene expression or activity. The database is useful for			
CC	producing electronic Northern blots that allow the user to determine the cell			
CC	type or tissue in which a given gene is expressed and to allow			
CC	determination of the abundance or expression level of a given gene in a			
CC	particular tissue or cell. The methods are useful for determining the			
CC	similarity of a toxic response to one or more individual compounds. The			
CC	methods are useful for predicting at least one toxic response or the			
CC	likelihood that a compound or test agent will induce various specific			
CC	pathologies such as those of the liver (liver necrosis, fatty liver			
CC	disease, protein adduct formation or hepatitis), those of the kidney,			
CC	heart, brain or testes, or other pathologies associated with at least one			
CC	of the toxins. The methods are also useful for predicting or elucidating			
CC	the potential cellular pathways influenced, induced or modulated by the			
CC	compound or test agent due to the similarity of the expression profile			
CC	compared to the profile induced by a known toxin. The present sequence is			
CC	that of a canine DNA sequence which was claimed for use during the			
XX	production of a canine microarray of the invention.			
XX	Sequence 706 BP; 208 A; 113 C; 158 G; 227 T; 0 U; 0 Other;			
XX	Query Match			
XX	Best Local Similarity 61.8%; Score 650.6; DB 13; Length 706;			
XX	Matches 671; Conservative 0; Mismatches 34; Indels 0; Gaps 0;			
QY	349 GTTCCAGAGGCTGCTGTTTCTATGCTTTTCAATTTCTATCGAGAAATGTTCACTCAGAG	408		
Db	1 GTTCCAGAGGCTGCTGTTTCTATGCTTTTCAATTTCTATCGAGAAATGTTCACTCAGAG	60		
QY	409 ATGTACAGTTTGTCTAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTT	468		
Db	61 ATGTACAGTTTGTCTAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTT	120		
QY	469 AATGCAATTTGAACACCTATGTTAAGAAAGGAAAGGAAATTTTATTT	528		
Db	121 AATGCAATTTGAACACCTATGTTAAGAAAGGAAAGGAAATTTTATTT	180		
QY	529 GCAGATAGAAATCTACTTTTGGGGAAGAGTGCTGCTGCTGTTAGAGGAGTT	588		
Db	181 GAAGATAGAAATCTACTTTTGGGGAAGAGTGCTGCTGCTGTTAGAGGAGTT	240		
QY	589 TTCTTCTCAGGATCTTTTGTGCTATATTTCTGCTTAAAGAGAGAGTCTTATGCCAGGA	648		
Db	241 TTCTTCTCAGGATCTTTTGTGCTATATTTCTGCTTAAAGAGAGAGTCTTATGCCAGGA	300		
QY	649 CTCACATTTTCCAAATGAACCTATCAGCAGAGATGAAGGAGTTTCACTGTGCTTGC	708		
Db	301 CTCACATTTTCCAAATGAACCTATCAGCAGAGATGAAGGAGTTTCACTGTGCTTGC	360		

RESULT 9

AAC78111

ID AAC78111 standard; cDNA; 1989 BP.

XX AAC78111;

AC AAC78111;

DT 08-FEB-2001 (first entry)

XX Human cancer associated gene sequence SEQ ID NO:505.

DE Human; cancer associated gene; cancer antigen; detection; cancer;

XX diagnosis; cytostatic; proliferative; vulnary; immunomodulator;

KW antiadipetic; antidiabetic; antihemorrhagic; antithrombotic; antitumor;

KW antiinflammatory; antithyroid; antiallergic; antidiabetic; antitumor;

KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;

KW vasotropic; antipsoriatic; antidiabetic; antitumor; immunomodulator;

KW immune disorder; haematopoietic cell disorder; autoimmune disorder;

KW allergic reaction; graft versus host disease; organ rejection;

KW haemostatic; thrombolytic; cardiovascular disorder; infection;

KW neurological disease; drug screening; ss.

XX Homo sapiens.

OS WO200055350-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US005882.

XX 12-MAR-1999; 99US-0124270P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2000-587533/55.

XX P-PSDB; AAB43902.

XX Novel isolated nucleic acids comprising sequences encoding peptides

XX useful for treating or diagnosing e.g. cancer.

XX Claim 1; Page 1035-1036; 2352pp; English.

XX AAC78111 to AAC78111 encode the human cancer associated proteins given in

XX AAB43902 to AAB44239. The proteins can have activities based on the

XX tissues and cells the genes are expressed in. Example of activities

XX include: cytostatic; proliferative; vulnary; immunomodulator;

CC anti-diabetic; antiasthmatic; antirheumatic; antiarthritic;
 CC anti-inflammatory; antihypertensive; antiallergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
 CC neoplastic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention
 XX
 SQ Sequence 1989 BP; 525 A; 439 C; 476 G; 544 T; 0 U; 5 Other;

Query Match 54.9%; Score 578.4; DB 3; Length 1989;
 Best Local Similarity 75.0%; Pred. No. 1.1e-154;
 Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCACTCTTAAGAAAGAGTTCTGCGCGTTTGTCTCTTCCAAATCCAGTACC 148
 DB 280 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTGTCTCTTCCCAATCCAGTACC 339

QY 149 CTGATATTGGAAATGTAACAGCAGCAGGCTTCTTCGACAGCAGAGAGGTGCG 208
 DB 340 ATGATATCTGGCAGATGTAAGAGGAGCAGGCTTCTTTCGACCGCGAGAGGTG 399

QY 209 ACTTATCAAGAGTCTCCCTCACTGCAACCAAGCTTAAAGCAGATGAGAAGTACTTCTCT 268
 DB 400 ACTCTCCAGGACATCTCAGCACTGGATCCCTGAAACCGAGGAGAGATATTTATAT 459

QY 269 CTCACATCTTAGCTTTTTCGACGAGTGAATGGAATGTAATGAAATTTTCGTGAGC 328
 DB 460 CCATGTTCTGGCTTCTTTCGACGAGCGATGGCATAGTAAATGAAATCTTGGTGGAGC 519

QY 329 GCTTTAGTCAGGAGTGCAGGCTTCCAGAGGCTGCTGTTTCTATGCGTTTCAAATTTCTCA 388
 DB 520 GATTTAGCCAAAGAGTTCAGATTACAGAGCCGCTGTTTCTATGCGTTTCAAATTTGCCA 579

QY 389 TCGAGAAATGTTCACTCAGAGATGTCAGATTTGCTGATAGACACTACATCAGAGATCCCA 448
 DB 580 TGAAGAAACATACATCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 639

QY 449 AGAAAGGGAAATTTTATTTAATGCAATTTGAAACCACTATGCTTAAAGAAAGAGCAG 508
 DB 640 AAGAAAGGGAATTTCTTCAATGCCATTTGAAACGATGCTTGTGTCAGAGAGAGCAG 699

QY 509 ATTGGCCCTTGGATGATAGCAGATAGAAATCTACTTTTGGGAAAGAGTGGTGCCT 568
 DB 700 ACTGGGCTTGGCTGATTTGGGGAACAAAGAGGCTACCTATGCTGAACGTGTGTAGCCT 759

QY 569 TTGCTGCTAGAGAGTTTCTTCTCAGGATCTTTCGCTGCTATATTTCTGCTTAAGA 628
 DB 760 TTGCTGCTAGAGAGCATTCTTCTTCCGCTTCTTTCGCTGATTTCTGCTCAAGA 819

QY 629 AGAGAGGCTTATGCGAGGACTCACCTTTTCCAAATGAACTCATCAGCAGAGATGAAGGAC 688
 DB 820 AACGAGGACTGATGCTGCGCTCACATTTCTTAATGACTTATTAGCAGATGAGGTT 879

QY 689 TTCATGTGACTTTGCTTCCCTGANGTTCCAAATCTAGTAAATAGCCTTCAAGAGAA 748
 DB 880 TACACTGTGATTTGCTTCCCTGATGTTCAACACACCTGGTACACAAACCATCGAGGAGA 939

QY 749 GGGTCAGGAGATCATTTGTGATGCTCTCAAAATTCAGCAGGAGTTTAAACAGAGCCT 808
 DB 940 GAGTAAAGAGAAATTAATATCAATGCTGTTTCGATAGAACAGGAGTTCCTCAGTGGCCT 999

QY 809 TGCCAGTTGGCCTCATTTGGAATGAAATTTGCAATTTTGATGAAACAGTACATTGAGTTGTAG 868

DB 1000 TGCCTGTGAAGCTCATTTGGGATGATTTGCATCTTAATGAGCAATACATTGATTTGTGG 1059
 QY 869 CTGACAGATTACTTTGTGGAACCTTGGATTTCTCAAAGTTTTCAGGCAGAAAATCCTTTTG 928
 DB 1060 CAGACAGACTTATGCTGGAACCTGGGTTTTCAGAGTTTTCAGACTAGAGAACCCATTG 1119
 QY 929 ATTTTATGGAACATTTCTTTAGAGGAAAAAACHAATTTCTTTGAGAAACAGATTTCAG 988
 DB 1120 ACTTTATGGAGATATTTTCACTGGAAGGAAAGACTTAACCTTCTTTGAGAGAGATGGCG 1179
 QY 989 AGTATCAGCGTTTTCAGTTATGTCAGAAACACAGATAACGCTTTCACCTTGGATGCGAG 1048
 DB 1180 AGTATCAGAGATGGAGTGAATGTCAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTG 1239
 QY 1049 ATTT 1052
 DB 1240 ACTT 1243

RESULT 10
 ADK70302
 ID ADK70302 standard; cDNA; 2215 BP.
 XX ADK70302;
 DT 06-MAY-2004 (first entry)
 XX
 DE Respiratory disease differentially expressed cDNA #38.
 XX
 KW ds; gene; cytostatic; respiratory; antiasthmatic; gene therapy;
 KW differential gene expression; respiratory disorder; lung cancer;
 KW chronic obstructive pulmonary disease; emphysema; asthma.
 XX
 OS Homo sapiens.
 XX
 PN WO2003101283-A2.
 XX
 PD 11-DEC-2003.
 XX
 PF 02-JUN-2003; 2003WO-US017409.
 XX
 PR 04-JUN-2002; 2002US-0386005P.
 XX
 PA (INCY-) INCYTE CORP.
 XX
 PI Rickert PK, Krasnow R;
 XX
 DR WPI; 2004-042945/04.
 XX
 PT New combination comprising cDNAs and proteins that are differentially
 PT expressed in respiratory disorders, useful for diagnosing or treating
 PT respiratory diseases e.g. lung cancer, chronic obstructive pulmonary
 PT diseases or asthma.
 XX
 PS Claim 1; SEQ ID NO 38; 343pp; English.
 XX
 CC The invention relates to cDNA sequences that are differentially expressed
 CC in respiratory disorders or their complements or encoded proteins. The
 CC cDNAs and proteins are useful for diagnosing, treating or monitoring
 CC treatment of a subject with a respiratory disease including lung cancer,
 CC chronic obstructive pulmonary diseases, emphysema or asthma. The protein
 CC is also useful for screening molecules or compounds to identify at least
 CC one ligand which specifically binds the protein. It is also useful for
 CC preparing and purifying a polyclonal or monoclonal antibody. This
 CC sequence corresponds to a cDNA of the invention.
 XX
 SQ Sequence 2215 BP; 593 A; 476 C; 523 G; 623 T; 0 U; 0 Other;

Query Match 54.9%; Score 578.4; DB 12; Length 2215;
 Best Local Similarity 75.0%; Pred. No. 1.2e-154;
 Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

Qy	89	ATGAAGAGCCATCTCTAAGAAAGAGTCTCGCGGTTTGTTCATCTTTCCATCCGATACC	148
Db	283	AGGATGAGCGCGTCTGAGAGAAAAACCCGCGCGCTTTGTCACTCTCCCATCGATACC	342
Qy	149	CTGATATTTGGAAAAATGATAAACAGGCAACAGGCTTCTTCTGACACAGAGAAGAGTGG	208
Db	343	ATGATATCTGGCAGATGTATAGAGCGCAGAGGCTTCCITTTGGACCGCGAGGAGTGG	402
Qy	209	ACTTATCAAGGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTACTTCATCT	268
Db	403	ACCTCTCCAAGGACATTCAGCACTTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT	462
Qy	269	CTCACATCTTAGGCTTTTTCGAGCCAGTGATGGAATGTAAATGAAAAATTTGGTGGAC	328
Db	463	CCCATGCTTCGGCTTTCTTTGCAAGCAAGCATGGCATAGTAATGAAAACTTGGTGGAC	522
Qy	329	GCTTTAGTCAGGAGTGCAGGTTCCAGAGGCTCGCTGTCTTCTATGGCTTTCAAATCTCA	388
Db	523	GATTTAGCCNAGAGTTCAGATTACAGAGCCGCTGTCTTATGGCTTCCAAATTTGCCA	582
Qy	389	TCGAGAATGTTCACTCAGAGATGTACAGTTGCTGATAGACACTTATCATCAGAGATCCCA	448
Db	583	TGGAATAACATACATCTCGAAATGTATAGTCTTCTTATTGACACTTATCATAAAAAGATCCCA	642
Qy	449	AGAAAAAGGAATTTTATTTAATGCAATTGAAACCATGCCCTATGTTTAGAAAAAGCAG	508
Db	643	AGAAAAAGGAATTTCTCTTCAAATGCCAATGAAACGATGCCCTGTGTCAAGAGAGCAG	702
Qy	509	ATTGGGCTTTGCCATGGATAGCAGATAGAAAACTACTTTTGGGGAAGAGTGGTGGCCT	568
Db	703	ACTGGGCTTTGCCATTTGGGACAAAGAGGCTACCTATGTGGAACGTGTGTAGCCT	762
Qy	569	TTGCTGCTAGAAAGGATTTTCTTCTCAGGATCTTTTGTGCTATATTTCTGGCTAAAGA	628
Db	763	TTGCTGCAGTGGAGGCAATTTCTTTTCCGGTCTTTTGTGCTGATATTTCTGGCTCAAGA	822
Qy	629	AGAGAGTCTTATGCCCCAGGACTCACATTTTCCAAATGAACCTCATCAGCAGAGATGAAGGAC	688
Db	823	AACGAGACTGTATGCCCTGGCCTCACATTTTCTAATGAACTTATAGCAGAGATGAGGTT	882
Qy	689	TTCACTGTGACTTTGCTTGCTGATGTTCCAAATCTTAGTAAATAAGCCTTCAGAAAGAA	748
Db	883	TACACTGTGATTTTGTCTTGCTGATGTTCAACACCTGGTACACAAACCATCGAGGAGA	942
Qy	749	GGGTCAAGGAGATCATTTGTATGCTCTCAAAATTTGAGCAGGAGTTTTTHAACAGAGCCT	808
Db	943	GAGTAAAGAGAAATAAATTTATCAATGTCTGTCGGATAGAACAGGAGTTCCTCACTGAGCCT	1002
Qy	809	TGCCAGTTGGCCTCATTTGGAATGAATTCGATTTTGATGAAACAGTACATTCAGTTGTAG	868
Db	1003	TGCCCTGTAGACTCATTTGGAGTGAATTTGCACTCTAATGAAGCAATACATTTGAGTTGTGG	1062
Qy	869	CTGACAGATTAATTGTGGAACTTTGGATTCCTCAAGTTTTCAGGCGAAGAAATCTCTTTTG	928
Db	1063	CAGACACATTAATGCTGGAATCGGTTTTTAGCAAGGTTTTTTCAGAGTAGAAACCCATTTG	1122
Qy	929	ATTTTATGGAACATTTCTTTAGAGGAAGAAACAAATTTCTTTTGAAACGAGTTTCAG	988
Db	1123	ACTTTATGGAGAATATTTTCACTGGAAGGAAGACTTAACCTTCTTTGAGAGAGATGAGCG	1182
Qy	989	AGTATCAGCGTTTTCAGAGTTATGGCAGAAACCAAGATAACGCTTCTCACCTTGGATGAC	1048
Db	1183	AGTATCAGAGGATGGGATGTATGTCAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTG	1242
Qy	1049	ATTT 1052	
Db	1243	ACTT 1246	

RESULT 11
ADJ56536
ID ADJ56536 standard; cDNA; 2216 BP.
XX

AC	AU56536;
XX	
DT	06-MAY-2004 (first entry)
XX	
DE	Murine cDNA differentially expressed in MYCN activated cells SeqID 342.
XX	
KW	mouse; murine; differential expression; transactivator; proto-oncogene;
KW	neuroblastoma; small cell lung cancer; cycostatic; gene therapy; ss;
KW	MYCN activated cell.
XX	
OS	Mus musculus.
XX	
FN	US2003119009-A1.
XX	
FD	26-JUN-2003.
XX	
PF	25-FEB-2002; 2002US-00084817.
XX	
PR	23-FEB-2001; 2001US-0270784P.
XX	
PA	(STUA/) STUART S G.
PA	(NUCH/) NUCHTERN J G.
PA	(PLON/) PLON S E.
PA	(SHOH/) SHOHEI J M.
XX	
PI	Stuart SG, Nuchtern JG, Plon SE, Shohet JM;
DR	WPI; 2003-635698/60.
XX	
PT	New genes regulated by MYCN activation, useful in gene therapy,
PT	particularly for treating a subject with e.g. neuroblastoma or other
PT	cancers, or for diagnosing, staging or monitoring the treatment of the
PT	cancer.
XX	
PS	Claim 1; SEQ ID NO 342; 27pp; English.
XX	
CC	This invention relates to novel isolated cDNAs that are differentially
CC	expressed in MYCN activated cells. Specifically, it refers to
CC	polynucleotide sequences that exhibit differential expression patterns in
CC	cells activated by the transactivator MYCN, where MYCN is a proto-
CC	oncogene that is amplified in neuroblastoma cells and is common in small
CC	cell lung cancers. The present invention describes these cDNA molecules
CC	as useful for in hybridisation assays to detect expression of nucleic
CC	acids (or complementary nucleic acids) in a present in a given sample, as
CC	well as for screening assays by identifying molecules or compounds that
CC	specifically bind the cDNA as a ligand and modulate function or activity.
CC	Accordingly, these compositions exhibit cytostatic activity and can also
CC	be used for gene therapy purposes. This polynucleotide sequence is a cDNA
CC	that is differentially expressed in MYCN activated cells, given in an
CC	exemplification of the invention. NOTE: This sequence does not appear in
CC	the printed specification but has been obtained in electronic format from
CC	the US Patent Office at
CC	ftp.seqdata.uspto.gov/sequence.html?DocID=20030119009.
XX	
SQ	Sequence 2216 BP; 593 A; 476 C; 524 G; 623 T; 0 U; 0 Other;
	Query Match 54.9%; Score 578.4; DB 10; Length 2216;
	Best Local Similarity 75.0%; Pred. No. 1.2e-154;
	Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;
QY	89 ATGAAGAGCCACTCCTTAAGAAGAGTTCTCGCCGGTTTGTTCATCTTCCAAATCCAGTACC 148
DB	284 AGGATGAGCGCTGCTGAGAGAAAACCCGCCGCTTTGTTCATCTTCCCCATCGAGTACC 343
QY	149 CTGATATTTCGAAATGTTATAAAGAGCAGCAGGCCTTCTTCTGACAGCAGCAAGAGGTGCG 208
DB	344 ATGATATCTGCGCATGTATTAAGAAGGACAGAGGCTTCTCTTTTGACCCGCCGAGGAGGTGG 403
QY	209 ACTTATCAAAGGATCTCCCTCACCTGGAAACAAGCTTTAAAGCAGATGAGAAGTACTTTCATCT 268
DB	404 ACCTCTCCAGGACATTCAGACATCGGGNATCCCTGAAACCCGAGGAGAGATATTTTATAT 463
QY	269 CTCATCTTAGCCCTTTTTTTCGACCCAGTGANGAAATTGTAATGAAAAATTTGGTGAGC 328

Db 1966 ACCTCTCCAGGACATTCAGCATCGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 1907
QY 269 CTCACATCTTAGCCTTTTTTGCAGCCAGTGAATTTGTAATGAAATTTGGTGGAGC 328
Db 1906 CCCATGTTCTGGCTTCTTTGACGACGATGGCATAGTAATGAAACTTTGGTGGAGC 1847
QY 329 GCTTTAGTCAGGAGGTCGAGGTTCCAGAGGCTGCCTGTTTCTATGGCTTTCAAATTTCTCA 388
Db 1846 GATTTAGCCAAAGAGTTCAGATTACAGAAAGCCGCTGTTTCTATGGCTTTCAAATTTGCCA 1787
QY 389 TCGAGATGTTCACTCAGAGATGTCAGTTTGCCTGATAGACATTCATCAGAGATGCCCA 448
Db 1786 TGGAAACATACATTCGAAATGATAGTCTTCTTATGACATTCATATAAAGATGCCCA 1727
QY 449 AGAAAAGGCAATTTTATTAATGCAATGAAACCATGCCCTATGTTTGAAGAAAAAGCAG 508
Db 1726 AAGAAAGGCAATTTCTTCAATGCCATTTGAACGATGCTTGTGTCAAGAAAGGCGAG 1667
QY 509 ATTGGCCCTTGGCATGGATAGCAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCCT 568
Db 1666 ACTGGGCCCTTGGCTGGATTTGGGGAACAAGAGGCTACTATGCTGAACGTTGTAGCCT 1607
QY 569 TTGCTCTGTAGNAGAGTTTCTTCTCAGGATCTTTTGGCTATATATTTGGCTTAAGA 628
Db 1606 TTGCTGCAGTGGAGGCAATTTCTTTCCGGTCTTTTGGCTCGATATTTGGCTCAAGA 1547
QY 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCAAATGAATCATCAGCAGAGATGAAGGAC 688
Db 1546 AAGCAGAGCTGATGCTGGCTCACATTTCTAATGAACTTATTAGCAGAGATGAGGTT 1487
QY 689 TTCATGTGCACTTTGCTTGGCTGATGTTCATPACTTTAGTAAATAGCCCTTCAGAGAAA 748
Db 1486 TACACTGTGATTTTGTCTGCTGATGTTTCAACACCTGGTACACAACCATCGAGGAGA 1427
QY 749 GGTTCAGGAGATCATTTGTCATGCTGTCAAAATTCAGCAGGAGTTTTCACAGAGCCT 808
Db 1426 GAGTAAGAGAAATAATATCAATGCTGTTTCGGATAGAACACGAGTTCCTCACTGAGCCT 1367
QY 809 TGCAGTTGCCCTCATTTGGGAATGAATTTGATGAAACAGTACATTCAGTTTGTAG 868
Db 1366 TGCTGTGAAGCTCATTTGGATGAATGCACTCTAATGAAGCAATACATTCAGTTTGTG 1307
QY 869 CTGACAGATTACTTTGGAACTTTGGAATTTCTCAAGGTTTTCAGGCGAGAAATCTTTTG 928
Db 1306 CAGCAGACTTATGCTGGACTGGTTTATGCAAGGTTTTCAGAGTAGAGAACCCATTG 1247
QY 929 ATTTATGAAAACATTTCTTTAGAGGAAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
Db 1246 ACTTTATGGAGATAATTTCACTGGAAGGAAAGACTTAATCTTTTGAAGAGAGTAGGCG 1187
QY 989 AGTATCAGCGTTTTCGAGTTTATGCGAAGAACCCAGATACGTTTCACCTTGGATGCG 1048
Db 1186 AGTATCAGAGGATGGGAGTGATGTCAGTCCACAGAGAAATCTTTTACCTTGGATGCTG 1127
QY 1049 ATTT 1052
Db 1126 ACTT 1123

RESULT 13
ABL65414
ID ABL65414 standard; DNA; 2500 BP.
XX AC
AC ABL65414;
XX
XX 15-MAY-2002 (first entry)
XX Lung cancer related gene sequence SEQ ID NO:3751.
DE Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
KW gene; ds.

XX Homo sapiens.
OS WO200194629-A2.
PN
XX
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
XX 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
PA (AVAL-) AVALON PHARM.
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI; 2002-198264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
XX Claim 1; SEQ ID NO 3751; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 16, 2005, 09:39:47 ; Search time 640 Seconds
(without alignments)
3246.608 Million cell updates/sec

Title: US-10-698-228-1
Perfect score: 1821
Sequence: 1 MGDPPEAAGLQDERS...QRFVMAETDNTVFTLDADF 351

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DRV=xlh
-Q=/cgn2_1/USFTO_spool/US10698228/runat_14062005_161415_28287/app_query.fasta_1.519
-DB=N Geneseq_16Dec04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228 @CGN_1_1_644 @runat_14062005_161415_28287 -NCPUL=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_16Dec04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	100.0	1053	4	Aaf32438 Human rib
2	1821	100.0	1053	4	Aaf32447 Human rib
3	1821	100.0	1081	4	Aaf32440 Human rib
4	1821	100.0	2596	3	Aa12411 cDNA enco
5	1821	100.0	4955	4	Aaf32439 Human rib

6	1821	100.0	4955	13	ADR24210
7	1817	99.8	1601	4	AaH14924
8	1460.5	80.2	1989	3	AaH78111
9	1460.5	80.2	2215	12	ADK70302
10	1460.5	80.2	2216	10	ADJ56536
11	1460.5	80.2	2482	4	AAS44917
12	1460.5	80.2	2500	6	ABL65414
13	1460.5	80.2	2500	6	ABL65517
14	1460.5	80.2	2500	6	ABL65859
15	1460.5	80.2	2500	8	ABX10335
16	1460.5	80.2	2500	11	ADI32044
17	1460.5	80.2	2500	12	ADN04443
18	1460.5	80.2	2500	12	ADQ09273
19	1460.5	80.2	2500	13	ACN37637
20	1460.5	80.2	3294	4	AAH73225
21	1460.5	80.2	3393	12	ADN03788
22	1460.5	80.2	3393	12	ADO19225
23	1447.5	79.5	2641	4	AAS44745
24	1429	78.5	1170	2	AAV05641
25	1425.5	78.3	1328	12	ADO57308
26	1376.5	75.6	1371	5	AAS79474
27	1348	74.0	1289	4	ABL14627
28	1241.5	68.2	3945	4	ABL14626
29	1230	67.5	1218	8	ABT20895
30	1220.5	67.0	1292	13	ADS49379
31	1210	66.4	14176	2	AAT84564
32	1210	66.4	14176	4	AAP84949
33	1210	66.4	14176	4	AAP84948
34	1192.5	65.5	1146	8	ABT19075
35	1192	65.5	1242	6	ABZ32250
36	1188	65.2	706	13	ADQ57092
37	1176.5	64.6	963	13	ADS58378
38	1168.5	64.2	1200	10	ACC61238
39	1168.5	64.2	1200	10	ADK63241
40	1168.5	64.2	1200	13	ADT47419
41	1161.5	63.8	1173	13	ADT47814
42	1161.5	63.8	1206	10	ADB69740
43	1151	63.2	1314	8	ABT18481
44	1151	63.2	1450	8	ABT20297
45	1151	63.2	3314	8	ABT17887

ALIGNMENTS

RESULT 1
AAF32438
ID AAF32438 standard; cDNA; 1053 BP.
XX
AC AAF32438;
XX
DT 18-APR-2001 (first entry)
XX
DE Human ribonucleotide reductase TP53R2H nucleotide sequence SEQ ID NO:2.
XX
KW Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
XX
OS Homo sapiens.
XX
FN WO200100799-A1.
XX
PD 04-JAN-2001.
XX
PF 27-JUN-2000; 2000WO-JP004189.
XX
PR 28-JUN-1999; 99JP-00181131.
PR 06-JUL-1999; 99JP-00192391.
PR 21-JAN-2000; 2000JP-00017770.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
PA (NAKA/) NAKAMURA Y.
PI Nakamura Y, Arakawa H, Tanaka H;
XX

DR WPI: 2001-112446/12.
 XX P-PSDB; AAB69050.

PT Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
 PT diagnosis, treatment and prevention of cancer.

XX Claim 5; Fig 1-3; 102pp; Japanese.

XX The present sequence encodes a human ribonucleotide reductase designated
 CC TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of the DNA
 CC repair mechanism and its activity is induced by p53. It can be used for
 CC the treatment, prevention and diagnosis of a wide range of cancers

XX SQ Sequence 1053 BP; 314 A; 192 C; 246 G; 301 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4,17e-212 Length: 1053
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-10-698-228-1 (1-351) x AAF32438 (1-1053)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20

DB 1 ATGGGCGACCCGGAAGCGGGAAGCGCGCGCTGGATCAGGATGAGAGATCATCTTCA 60

QY 21 AspThrAsnGluSerGluLeuIleValSerAsnGluGluProLeuLeuArgLysSerArg 40

DB 61 GACACCAACGAAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 120

QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60

DB 121 CGGTGTGTCATCTTCCATCCAGTACCTGATATTGGAAATGATATAACAGGCACAG 180

QY 61 AlaSerPheTyrThrAlaGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80

DB 181 GCTTCCTCTGGACAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCAGTGAACAG 240

QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaIleSerAsp 100

DB 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCAGATCTTACGCTTTTTCAGCAGCATGAT 300

QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120

DB 301 GGAATGTAAATGAATTTTGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360

QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140

DB 361 CGCTGTTTCTATGGCTTTCAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTG 420

QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160

DB 421 CTGATAGACACTTACATCAGATCCCAAGAAAGGAAATTTTATTTAATGCAATGAA 480

QY 161 ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180

DB 481 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGGCTTGGATGGATGATAGAGAA 540

QY 181 SerThrPheGlyGluArgValAlaAlaPheAlaAlaValGluGlyValPhePheSerGly 200

DB 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTAGAGAGAGTTTCTTCTCAGGA 600

QY 201 SerPheAlaAlaIlePheTyrLysLysArgGlyLeuMetProGlyLeuThrPheSer 220

DB 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTTCTTATGCCAGGACTCACTTTTCC 660

QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240

DB 661 AATGAACATCATCAGCAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720

QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260

DB 721 TACTTAGTAATTAAGCTTTCAGAGAAAGGGTCAGGAGAGATCATTTGTTGCTGTCAA 780

QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280

DB 781 ATTGAGCAGGAGTTTAAACAGAAAGCTTCCAGTTCGCCTCATTTGAATGAATTCATT 840

QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300

DB 841 TTGATGAACACAGTACATTTGAGTTTGTAGCTGACAGATTTACTTGTGAACCTTGGATTCTCA 900

QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320

DB 901 AAGGTTTTTCAGGAGAAATCTTTTGTATTTATGAAACATTTCTTTAGAGAGAAA 960

QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340

DB 961 ACAATTTCTTTCAGAAACAGGTTTTCAGATATCAGGTTTTCAGATTATGCGAGAAACC 1020

QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

DB 1021 ACAGATAACGCTTTCACCTTGGATGCGAGATTTT 1053

RESULT 2

AAF32447

ID AAF32447 standard; cDNA; 1053 BP.

XX AAF32447;

XX 18-APR-2001 (first entry)

XX Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:12.

XX Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.

XX Homo sapiens.

XX WO200100799-A1.

XX 04-JAN-2001.

XX 27-JUN-2000; 2000WO-JP004189.

XX 28-JUN-1999; 99JP-00181131.

XX 06-JUL-1999; 99JP-00192391.

XX 21-JAN-2000; 2000JP-00017770.

XX (TAKE) TAKEDA CHEM IND LTD.

XX (NAKA/) NAKAMURA Y.

XX Nakamura Y, Arakawa H, Tanaka H;

XX WPI; 2001-112446/12.

XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for

XX diagnosis, treatment and prevention of cancer.

XX Claim 6; Page 96-97; 102pp; Japanese.

XX The present invention describes a human ribonucleotide reductase

XX designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of

XX the DNA repair mechanism and its activity is induced by p53. It can be

XX used for the treatment, prevention and diagnosis of a wide range of

XX cancers. The present sequence represents a human ribonucleotide reductase

XX related sequence which is given in the exemplification of the present

XX invention

XX SQ Sequence 1053 BP; 314 A; 191 C; 246 G; 302 T; 0 U; 0 Other;

XX Alignment Scores:

XX Pred. No.: 4,17e-212 Length: 1053

XX Score: 1821.00 Matches: 351

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-10-698-228-1 (1-351) x AAF32447 (1-1053)

Qy	1	MetGlyAspProGluuArgProGluuAlaAlaGlyLeuAspGluAspGluuArgSerSer	20
Db	1	ATGGCGCACCCGGAAGACCGGAGCGCGCGCTGGATCAGGATGAGAGATCATCTTCA	60
Qy	21	AspThrAsnGluSerGluileLysSerAsnGluGluProLeuLeuArgLysSerSerArg	40
Db	61	GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCCTAAGAAGAAGTTCCTGC	120
Qy	41	ArgPheValIlePheProIleGlnTyrProAspIleTTrpLysMetTyrLysGlnAlaGln	60
Db	121	CGGTTTGTCATCTTTCCAATCAGTACCCTGATATTTGGAAATATGATATAACAGGACACAG	180
Qy	61	AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys	80
Db	181	GCCTTCCTCTGGACAGCAAGAGGTGACTTATCAAGGATCTCCTCACTGGAAACAAG	240
Qy	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaIaSerAsp	100
Db	241	CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTTAGCCCTTTTTCAGCCAGTGAT	300
Qy	101	GlyIleValAsnGluAsnLeuValGluuArgPheSerGlnGluValGlnValProGluuAla	120
Db	301	GGAAATGTAAATGAAATTTTGGTGAGGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT	360
Qy	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluuMetTyrSerLeu	140
Db	361	CGCTGTTTCTATGGCTTTCAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG	420
Qy	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160
Db	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA	480
Qy	161	ThrMetProTyrValLysLysValaAspTrpAlaLeuArgTrpIleAlaAspArgLys	180
Db	481	ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTTGCGATGGATAGCAGATAGAAA	540
Qy	181	SerThrPheGlyGluuArgValValaPheAlaAlaValGluGluValPhePheSerGly	200
Db	541	TCTACTTTTGGGAAAGAGTGGTGCCCTTGCTGCTGTAGAAAGGAGTTTCTTCTCAGGA	600
Qy	201	SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer	220
Db	601	TCTTTTGCTGCTATATTCTGGCTAAGAAGAGAGGTCTTATGCCAGGACTCACTTTTTC	660
Qy	221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln	240
Db	661	AATGAACTCATCAGCAGAGATGAAGGACTTCACCTGTGACTTTGCTGCTGATGTTCCAA	720
Qy	241	TyrLeuValAsnLysProSerGluGluuArgValArgGluIleIleValAspAlaValLys	260
Db	721	TACTTTAGTAAATAAGCCCTTCAGAAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAA	780
Qy	261	IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle	280
Db	781	ATTGAGCAGGAGTTTTTAAACAGAGCCCTTGCAGTTGGCTCATTTGGGAATGAATTCATT	840
Qy	281	LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer	300
Db	841	TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTTGTGGAACTTGGATTCTCA	900
Qy	301	LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys	320
Db	901	AAGGTTTTTTCAGGCAGAAAATCCCTTTGATTTTATGGAAAAATTTCTTTTAGAAGGAAA	960
Qy	321	ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr	340

Db 961 ACAAAATTCCTTTGGAAACAGATTTTCAGAGTATCAGCGCTTTTCAGTTATGGCAGAAACC 1020
 Qy 341 ThrAspAenValPheThrLeuAspAlaAspPhe 351
 Db 1021 ACAGATAACGCTCTTCACCTTGGATGCAGATTTT 1053

RESULT 3

AAF32440
ID AAF32440 standard; cDNA: 1081 BP.

AC AAF32440:

DT 18-APR-2001 (first entry)

XX DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:4.

Human: ribonucleotide reductase: cancer: DNA repair: p53: ss:

XX
OS Homo sapiens.

XX
PN · WO200100799-A1.XX
PD
04-JAN-2001

XX 27-JUN-2000: 2000WO-JP004189.

XX
PR 28-JUN-1999: 99JP-00181131.

PR 06-JUL-1999; 99JP-00192391.
PR 21-JAN-2000: 2000JP-00017770.

XX
PA (TAKE) TAKEDA CHEM IND LTD.PA (NAKA/) NAKAMURA Y.
XX

PI Nakamura Y, Arakawa H, Tanaka H, XX

DR WPI; 2001-112446/12.
XX

pt Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
pt diagnosis, treatment and prevention of cancer.

XX
PS Example 2: Page 91: 102pp: Japanese.

xx The present invention describes a human ribonucleotide reductase
cc designated TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of
cc the DNA repair mechanism and its activity is induced by p53. It can be
cc used for the treatment, prevention and diagnosis of a wide range of
cc cancers. The present sequence represents a human ribonucleotide reductase
cc related sequence which is used in an example from the present invention

Sequence 1081 BP: 322 A: 199 C: 252 G: 308 T: 0 U: 0 Other: 0

Alignment Scores:

Pred. No.:	4.34e-212	Length:	1081
Score:	1821.00	Matches:	351

Percent Similarity: 100.00%
Best [local] Similarity: 100.00%
Mismatch: 0.00%
Conservation: 0.00%

```

Query Match:
Indels:
Gaps:
4
100.00%

```

US-10-698-228-1 (1-351) x AAF32440 (1-1081)

1 MetGlyAspProGluArgProGluAlaAlaGly

db
20 ATGGGCGACCCGGAAGGCCGGAAGCGCCGGC

21 AspThrAsnGluSerGluIleLysSerAsnGlu

80 GACACCAACGAAAGTGAAATAAAGTCAAAATGAA

41 ArgpHeValIle

HeProIleGlnTyrProAsn

140 CCGTTTGTGCATCTTTCCAAATCCAGTACCCCTGAT

QY 61 AlaSerPheTrpThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 Db 200 GCTTCCTCTGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGACAAG 259
 QY 81 LeuLysAlaAspGluLysTrpPheIleSerHisIleLeuAlaPhePheAlaLaserAsp 100
 Db 260 CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 319
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 Db 320 CGAATTTGTAAGTAAATTTTGTGGAGCGCTTTAGTCAGAGGTGCAGGTCAGAGGCT 379
 QY 121 ArgCysPheTrpGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTrpSerLeu 140
 Db 380 CGCTGTTTCTATGGCTTTCAAATTTCTATCGAAGATTTTCACTCAGAGATGTACAGTTTG 439
 QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 Db 440 CTGATAGACATTTACATCAGATCCCAAGAAAGGGAATTTTATTTAATGCATTTGAA 499
 QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 Db 500 ACCATGCCCTATTTAAGAAAAGCAGATTTGGCCCTTCGGATGGATAGCAGATAGAAA 559
 QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 Db 560 TCTACTTTTGGGCAAGAGTGGTGGCTTTGCTGTGTAGAAAGGAGCTTTCTTCTCAGGA 619
 QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTrpPheSer 220
 Db 620 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 679
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db 680 ATGAACTCATCAGCAGATGAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 739
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 Db 740 TACTTTAGTAAATAAGCTTTCAAGAAAGGGTCCAGGAGATCATTTGTCATGCTGTCAAA 799
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 Db 800 ATTGAGCAGAGTGTATTAACAGAAAGCTTGGCCAGTTGGCTCATTTGGAAATGAATTCAT 859
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 Db 860 TTGATGAACAGTACTATGAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTTCTCA 919
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 Db 920 AAGGTTTTTTCAGGCAGAAAATCCTTTTGTATTTATGGAACATTTCTTTAGAAGGAAA 979
 QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 Db 980 ACAAAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGCTTATGGCAGAAAC 1039
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1040 ACAGATAACGTCTTACCTTGGATGAGATTTT 1072

RESULT 4

ID AAA12411 standard; cDNA; 2596 BP.

XX AAA12411;

XX 25-JUL-2000 (first entry)

XX cDNA encoding a human RNA-associated protein.

XX Human; RNA-associated protein; cell proliferation; cancer; inflammation;
 KW immune response; reproductive disorder; actinic keratosis;
 KW atherosclerosis; arteriosclerosis; bursitis; cirrhosis; hepatitis;

KW mixed connective tissue disease; myelofibrosis; primary thrombocythemia;
 KW paroxysmal nocturnal hemoglobinuria; polycythemia vera; psoriasis;
 XX trauma; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 56..1111
 FT /*tag= a
 FT /product= "RNA-associated protein"

XX W0200015799-A2.

XX 23-MAR-2000.

XX 17-SEP-1999; 99WO-US021688.

XX 17-SEP-1998; 98US-00156039.

XX 22-SEP-1998; 98US-00158720.

XX 04-NOV-1998; 98US-00186815.

XX 08-APR-1999; 99US-0128660P.

XX (INCY-) INCYTE PHARM INC.

XX Tang YT, Corley NC, Guegler KJ, Gorgone GA, Patterson C;
 PI Hillman JL, Baughn MR, Lal P, Azimzai Y, Yue H, Yang J;

XX WPI; 2000-271437/23.

XX P-PSDB; AAY84439.

XX New polypeptides and polynucleotides, useful for preventing and treating

XX a disorder associated with increased or decreased expression of RNA

XX associated proteins.

XX Claim 9; Page 120-121; 131pp; English.

XX The present sequence encodes a human RNA-associated protein. The
 CC expression of RNA-associated proteins is closely associated with
 CC reproductive tissues, nervous tissues, cell proliferation including
 CC cancer, inflammation and immune responses, and so they may be used for
 CC diagnosis, treatment or prevention of cell proliferative, diseases and
 CC immune/inflammatory disorders, and reproductive disorders. Diseases and
 CC disorders which may be treated include actinic keratosis,
 CC atherosclerosis, arteriosclerosis, bursitis, cirrhosis, hepatitis, mixed
 CC connective tissue disease, myelofibrosis, paroxysmal nocturnal
 CC hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia
 CC and cancers, and trauma

XX SQ Sequence 2596 BP; 783 A; 418 C; 510 G; 885 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1 59e-211 Length: 2596
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0

US-10-698-228-1 (1-351) x AAA12411 (1-2596)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnArgSerSer 20
 Db 56 ATGGCGACCCGGAAGAGCGCGGCGGCTGGATCAGGATCAGATCATCTTCA 115
 QY 21 AspThrAsnGluSerGluIleLysSerAsnGluPProLeuLeuArgLysSerSerArg 40
 Db 116 GACACCAACCAAGTGAATAAAGTCAATCAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 175
 QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
 Db 176 CGGTGTGTCTATCTTCCAAATCCAGTACCTTGATATTTGGAAATATGATAAACAGGCACAG 235
 QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80

```

Db      236  GCTTCTCTTCTGGACACAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAAACAAG 295
Qy      81  LeuLeuAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db      296  CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTTACGCTTTTTCGAGCAGTGAT 355
Qy      101  GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db      356  GGAATTGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGCT 415
Qy      121  ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db      416  CGCTGTTCTATGGCTTTCAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 475
Qy      141  LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db      476  CTGATAGACACTTACATCAGATGCCAGATCCCAAGAAAGGAATTTTATTAATGCAATGAA 535
Qy      161  ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
Db      536  ACCATGCCCTATGTTAAGAAAGAAAGCAGATTGGGCTTGGATGGATAGCAGATAGAAA 595
Qy      181  SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 200
Db      596  TCTACTTTTGGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGTTTCTTCTCAGGA 655
Qy      201  SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db      656  TCTTTTGTCTATATCTGGCTTAAGAGAGAGGCTTATGCCAGGACTCACTTTTTC 715
Qy      221  AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db      716  AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 775
Qy      241  TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db      776  TACTTAGTAATAAGCCTTCAGAAAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 835
Qy      261  IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db      836  ATTGACACAGAGTTTTTAACAGAAAGCCTTGCCAGTTGGCCTCATTTGGAATGGAATTC 895
Qy      281  LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db      896  TTGATGAAACAGTACATTGAGTTTGTAGCTCAGAGATTACTTGTGGAACCTTGGATTCTCA 955
Qy      301  LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db      956  AAGGTTTTTTCAGGCAGAAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAAGGAAA 1015
Qy      321  ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db      1016  ACAAAATTTCTTTGAGAAACAGGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1075
Qy      341  ThrAsnValPheThrLeuAspAlaAspPhe 351
Db      1076  ACAGATAACGCTTTCACCTTGGATGAGATTTT 1108

RESULT 5
AAF32439
ID  AAF32439 standard; cDNA; 4955 BP.
XX
AC  AAF32439;
XX
DT  18-APR-2001 (first entry)
XX
DE  Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:3.
XX
KW  Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
XX
OS  Homo sapiens.
XX

```

```

PN      WO200100799-A1.
FD      04-JAN-2001.
XX      27-JUN-2000; 2000WO-JP004189.
XX      28-JUN-1999; 99JP-00181131.
PR      06-JUL-1999; 99JP-00192391.
PR      21-JAN-2000; 2000JP-00017770.
XX      (TAKES) TAKEDA CHEM IND LTD.
PA      (NAKA/) NAKAMURA Y.
XX      Nakamura Y, Arakawa H, Tanaka H;
XX      WPI; 2001-112446/12.
XX      Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
PT      diagnosis, treatment and prevention of cancer.
XX      Example 2; Page 87-90; 102pp; Japanese.
XX      The present invention describes a human ribonucleotide reductase
CC      designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of
CC      the DNA repair mechanism and its activity is induced by p53. It can be
CC      used for the treatment, prevention and diagnosis of a wide range of
CC      cancers. The present sequence represents a human ribonucleotide reductase
CC      related sequence which is used in an example from the present invention
XX      Sequence 4955 BP; 1526 A; 803 C; 999 G; 1627 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      4.13e-211      Length:      4955
Score:          1821.00      Matches:      351
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              4              Gaps:      0

US-10-698-228-1 (1-351) x AAF32439 (1-4955)
Qy      1  MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db      245  ATGGCGCACCGGAAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
Qy      21  AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
Db      305  GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 364
Qy      41  ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
Db      365  CGGTTTGTATCTTTCATCCATCAGTACCTCATATTTGAAATAATGTATTAACAGGCACAG 424
Qy      61  AlaSerPheThrThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db      425  GCTTCTCTTGGACAGCAGAGAGTGCAGCTTATCAAGAGATCTCCCTCACTGGAACAAG 484
Qy      81  LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db      485  CTTAAAGCAGATGAGAAGTACTTCACTCTCTCACATCTTAGCCCTTTTTCGAGCAGTGAT 544
Qy      101  GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db      545  GGAATTGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGCT 604
Qy      121  ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db      605  CGCTGTTCTATGGCTTTCAATTTCTCATCGAATGTTCTACTCAGAGATGTACAGTTG 664
Qy      141  LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db      665  CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 724

```

QY 161 ThrMetProTyrVallylsLysAlaAspTrrAlaLeuArgTrrileAlaAspArgLys 180
 Db 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCCCTTCGATGGATAGCAGATAGAAA 784
 QY 181 SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
 Db 785 TCTACTTTGGGAAAGAGTGGTGGCTTTCGCTGTAGAGAGATTTCTTCTCAGGA 844
 QY 201 SerPheAlaAlailePheTrrPheLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 Db 845 TCTTTTGTCTATATCTGGCTAAAGAGAGAGAGGCTTATGCGCAGGACTCACTTTTCC 904
 QY 221 AnGluLeuLeuSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db 905 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluLeuValleValAspAlaValLys 260
 Db 965 TACTTAGTAAATAGGCTTCAGNAGAAAGGTCAGGAGATCAATTGTTGATGCTGTCAAA 1024
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuileGlyMetAsnCysile 280
 Db 1025 ATTGAGCAGGAGTTTAAACAGAGGCTTGCAGTTGGCTCATTTGGAATGAATTGCATT 1084
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 Db 1085 TTGATGAAACAGTACATTGAGTTTGTAGCTGCAGAGATTACTTGTGAACTTGGATTCTCA 1144
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnileSerLeuGluGlyLys 320
 Db 1145 AAGGTTTTTCAGGAGAGAAATCCTTTTGATTTTATGGAAACATTTCTTTAGAGGAGAAA 1204
 QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 Db 1205 ACAATTTCTTTGAGBAAACGATTTTCAGAGTATCAGCGTTTGCAGTTATGGCAGAAACC 1264
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1265 ACAGATAAGCTTCTACCTTGGATGCAGATTTT 1297

RESULT 6
 ADR24210
 ID ADR24210 standard; DNA; 4955 BP.
 XX AC ADR24210;
 XX DT 21-OCT-2004 (first entry)
 XX DE Breast cancer prognosis marker #71.
 XX KW ds; breast cancer; prognosis; gene expression; diagnosis.
 XX OS Homo sapiens.
 XX PN WO2004065545-A2.
 XX PD 05-AUG-2004.
 XX PF 15-JAN-2004; 2004WO-US0001100.
 XX PR 15-JAN-2003; 2003US-00342887.
 XX PA (ROSE-) ROSETTA INPHARMATICS LLC.
 XX PA (NECA-) NETHERLANDS CANCER INST.
 XX Van't Veer LJ, He Y;
 XX WPI; 2004-593473/57.
 XX Classifying a breast cancer patient according to prognosis comprises
 PT determining the similarity between the level of expression of each of
 PT five genes in a cell sample taken from patient, to control levels.
 XX

PS Disclosure; SEQ ID NO 71; 226pp; English.
 XX The invention relates to a method of classifying a breast cancer patient
 CC according to prognosis by determining the similarity between the level of
 CC expression of each of five genes for which markers are listed in the
 CC specification, in a cell sample taken from the breast cancer patient, to
 CC control levels of expression for each respective five genes to obtain a
 CC patient similarity value. The methods are useful for classifying a breast
 CC cancer patient according to prognosis. Kits and computer program products
 CC are useful for data analysis using the diagnostic, prognostic and
 CC statistical methods of the invention. This sequence corresponds to a
 CC marker used in the method of the invention.
 XX Sequence 4955 BP; 1526 A; 803 C; 999 G; 1627 T; 0 U; 0 Other;
 SQ

Alignment Scores:
 Pred. No.: 4,13e-211 Length: 4955
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 13 Gaps: 0

US-10-698-228-1 (1-351) x ADR24210 (1-4955)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
 Db 245 ATGGCGCACCGGAAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
 QY 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerArg 40
 Db 305 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCAAGAAAGATTCTCC 364
 QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrrLysMetTrrLysGlnAlaGln 60
 Db 365 CGTTTGTGATCTTTCATCCAGTACCTGATATTGGAAATGTATTAACAGGCACAG 424
 QY 61 AlaSerPheTrrPheAlaGluGluValAspLeuSerLysAspLeuProHisTrrPheLys 80
 Db 425 GCTTCTCTTCGGACAGCAGAGAGGTGACATTATCAAGAGGATCTCCCTCACTGGAACAAG 484
 QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAsp 100
 Db 485 CTTAAAGCAGATGAGAAGTACTTCTCTCCTCAGATCTTTAGCCCTTTTTCAGCCAGTGAT 544
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 Db 545 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTCCAGGCTCCAGAGCT 604
 QY 121 ArgCysPheTrrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTrrSerLeu 140
 Db 605 CGCTGTTTCTATGCTTTCAAAATTTCTATCGAGAATGTTCACTCAGAGATGTACAGTTTG 664
 QY 141 LeuIleAspThrTrrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTAATGCAATGAA 724
 QY 161 ThrMetProTrrVallylsLysAlaAspTrrAlaLeuArgTrrileAlaAspArgLys 180
 Db 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTGGCATGGATGAGATGAGAAAA 784
 QY 181 SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
 Db 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTCGCTGTAGAGAGATTTTCTTCTCAGGA 844
 QY 201 SerPheAlaAlailePheTrrPheLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 Db 845 TCTTTTGTCTATATCTGGCTAAAGAGAGAGGCTTATGCGCAGGACTCACTTTTCC 904
 QY 221 AsnGluLeuLeuSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db 905 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964

Qy 241 TyrluValAsnLysProSerGluGluArgValArgGluIleLeuValAspAlaValLys 260
 Db 965 TACTTAGTAATAAGCCCTTCAGAGAAAGGGTCAGGGAGATCATTTGTGCTGTCAAA 1024
 Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 Db 1025 ATTGAGCAGGAGTTTAAACAGAGCCTTGCCAGTTGGCCTCATTTGAATGAATTGAT 1084
 Qy 281 LeuMetLysGlnTyrlleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 Db 1085 TTGATGAACAGTACATGATGTTGTAGCTGACAGATTACTTTGGGAACCTGGATTCTCA 1144
 Qy 301 LysValPheGlnAlaGluAanProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 Db 1145 AAGTTTTTTCAGGAGAGAAATCCCTTTGATTTTATGGAAACATTTCTTTAGAGAGGAAA 1204
 Qy 321 ThrAsnPheGluLysArgValSerGluTyrglnArgPheAlaValMetAlaGluThr 340
 Db 1205 ACAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGGTTTTCAGATTATGGCAGAAACC 1264
 Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1265 ACAGATAACCTCTTCACCTTGGATGCAGATTTT 1297

RESULT 7

AAH14924

ID AAH14924 standard; cDNA; 1601 BP.

AC AAH14924;

XX 26-JUN-2001 (first entry)

XX Human cDNA sequence SEQ ID NO:12810.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EF1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.

XX 27-AUG-1999; 99JP-00300253.

XX 11-JAN-2000; 2000JP-00118776.

XX 02-MAY-2000; 2000JP-00183767.

XX 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesising polynucleotides, particularly the 5602 full-

XX length cDNAs defined in the specification, and for the detection and/or

XX diagnosis of the abnormality of the proteins encoded by the full-length

XX cDNAs.

XX Claim 8; SEQ ID NO 12810; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesising 5602 full-

XX length cDNAs defined in the specification. Where a primer set comprises:

XX (a) an oligo-dT primer and an oligonucleotide complementary to the

XX complementary strand of a polynucleotide which comprises one of the 5602

XX nucleotide sequences defined in the specification, where the

XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination

XX of an oligonucleotide comprising a sequence complementary to the

XX complementary strand of a polynucleotide which comprises a 5'-end

XX sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention

XX SQ Sequence 1601 BP; 484 A; 290 C; 348 G; 479 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,39e-211 Length: 1601
 Score: 1817.00 Matches: 350
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.72% Mismatches: 0
 Query Match: 99.78% Indels: 0
 DB: 4 Gaps: 0

US-10-698-228-1 (1-351) x AAH14924 (1-1601)

Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20

Db 40 ATGGCGGACCCGGAAGGCGGAGCGCGGGCTGGATCAGATGAGATCATCTTCA 99

Qy 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40

Db 100 GACACCAACGAAAGTGAATAAAGTCAAAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 159

Qy 41 ArgPheValIlePheProIleGlnTyrlProAspIleTyrlProLysMetTyrlLysGlnAlaGln 60

Db 160 CGGTTTGTTCATCTTTCCAAATCAGTACCTCATATTTGAAAATGTATATAACAGGCACAG 219

Qy 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80

Db 220 GCTTCCTTCTGGACAGCAGAGAGGTGCATTTATCAAGAGATCTCCCTCATCTGGAAACAAG 279

Qy 81 LeuLysAlaAspGluLysTyrlPheIleSerHisIleLeuAlaPhePheAlaAsp 100

Db 280 CTTHAAGCAGATGAGAGTACTTTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT 339

Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120

Db 340 GGAATTGTAAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 399

Qy 121 ArgCysPheTyrlGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrlSerLeu 140

Db 400 CGCTGCTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTCACTCAGAGATGTACAGTTG 459

Qy 141 LeuIleAspThrTyrlleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160

Db 460 CTGATAGACATTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATCAATTGAA 519

Qy 161 ThrMetProTyrlValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180

Db 520 ACCATGCCCTATTTAGAAAAAGCAGATTTGGGCTTTGGCATGTAGTAGAGATAAAA 579

Qy 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluValPhePheSerGly 200

Db 580 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGAGGAGTTTCTTCTCAGGA 639

Qy 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220

Db 640 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCATCTTTTC 699

Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240

Db 700 AATGAACCTCATCAGCAGAGATGAAGGACTTTCACCTGTGACTTTGCTTGTGCTGTGCTCA 759

QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260
 DB 760 TACTTAGTAAATAAGCCTTCAGAAGAAGGGTCAGGAGATCATTTGATGCTGTCAAA 819
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAnCysIle 280
 DB 820 NTGAGCAGAGATTTTAAACAGAGCCTTGCCAGTTGGCTCATTTGGAGTGAATTCAT 879
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 DB 880 TTGATGAACAGTACATTCAGTTGTAGCTGACAGATTAATCTGTGGAATTCGATTCCTCA 939
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 DB 940 AAGGTTTTTCAGGCAGAAAATCCTTTGATTTTATGGAATAACATTTCTTTAGAGGAAAA 999
 QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 DB 1000 ACAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1059
 QY 341 ThrAsnValPheThrLeuAspAlaAspPhe 351
 DB 1060 ACAGATAACGTCTTCACCTTGGATGCAGATTTT 1092
 RESULT 8
 AAC78111
 ID AAC78111 standard; cDNA; 1989 BP.
 XX
 AC AAC78111;
 XT
 DT 08-FEB-2001 (first entry)
 XX
 DE Human cancer associated gene sequence SEQ ID NO:505.
 XX
 KW Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
 KW antidiabetic; antiaslathmic; antirheumatic; antiarthritic; antiviral;
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
 KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
 KW vasotropic; antiporiatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO20005350-A1.
 XX
 XX 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US005882.
 XX
 PR 12-MAR-1999; 99US-0124270P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 XX WPI; 2000-587533/55.
 DR P-PSDB; AAB43902.
 XX
 PT Novel isolated nucleic acids comprising sequences encoding peptides
 useful for treating or diagnosing e.g. cancer.
 XX
 PS Claim 1; Page 1035-1036; 2352pp; English.
 XX
 CC AAC77607 to AAC78448 encode the human cancer associated proteins given in
 CC AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnery; immunomodulator;
 CC antidiabetic; antiaslathmic; antirheumatic; antiarthritic;

CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
 CC vasotropic; antiporiatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies, and
 CC agonists and antagonists may be also used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention
 XX
 SQ Sequence 1989 BP; 525 A; 439 C; 476 G; 544 T; 0 U; 5 Other;

Alignment Scores:
 Pred. No.: 1,21e-167 Length: 1989
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 3 Gaps: 1

US-10-698-228-1 (1-351) x AAC78111 (1-1989)

QY 4 ProGluArgProGluAlaAlaGlyLeuAspGluAspGluArgSerSerSerAspThrAs 23
 DB 185 CCGAGCGGAGCCCGGCTCTGCGCCAGCAAGACCGCGAGGAGATCTTCCAGGAGCCAC 244
 QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
 DB 245 GAGCGGAAACTAAGCAGCTGCCCGCGGTGGAGATGAGCGCTGTGAGAGAAA 304
 QY 38 rSerArgPheValIlePhePheProIleGlnTyrProAspIleTyrLysG 58
 DB 305 CCGCGCGCGCTTGTCTATCTCCCATCGAGTACCATGATCTGGCAGATGTATAAGAA 364
 QY 58 nAlaGlnAlaSerPheThrThraGluGluValAspLysSerLysAspLeuProHisTr 78
 DB 365 GCGAGCGCTTCTTTTGGACCCCGAGAGGTGGACCTCTCCAGGACATTCAGACTG 424
 QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaA 98
 DB 425 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCATGTTCTGGCTTCTTTCAGC 484
 QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
 DB 485 AAGCGATGGCATAGTAAATGAAACTTGTGGAGCGATTTAGCCAGAGTTTCAGATTAC 544
 QY 118 oGluAlaArgCysPheThrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
 DB 545 AGAAGCGCGCTTCTTCTATGGCTTCCAAATTTGCCAATGCGAATAACATATCTGAAATGA 604
 QY 138 rSerLeuLeuLeuAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
 DB 605 TAGTCTTCTTATTGACACTTACATAAAGATCCCAAGAAAGGGAATTTCTTCAATGC 664
 QY 158 alIeGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
 DB 665 CATTAAGACGATGCTTGTGTCAAGAGAGGAGACTGGGCTTGGCTGGATGGGA 724
 QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
 DB 725 CAAAGAGGCTACCTATGCTGAACGTTGTAGCTTTGCTGCTGAGTGAAGCAATTTCTT 784
 QY 198 eSerGlySerPheAlaAlaIlePheThrPheLysLysArgGlyLeuMetProGlyLeuTh 218
 DB 785 TTCGGTCTCTTTTGGCTCGATATTTGGCTCAAGAAACGAGGAGTGTGCTGCCTCAC 844

QY 218 rPheSerAsnGluLeuIleSerArgPheGluGluGluHisCysAspPheAlaCysLeuMe 238
 DB 845 ATTTTCTTAATGAACCTTATTACGAGATGAGGGTTTACACTGTGATTTGCTTGCCTGAT 904
 QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluGluArgValArgGluIleValAspAl 258
 DB 905 GTTCMAACACCTGGTACACAAACCATCGGAGGAGAGTAAGAGAAATAATTATCAATGC 964
 QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 DB 965 TGTTCGGATAGAACAGGAGTTCCTCACTGAGGCCTTGCCTGTGAAGCTCATTTGGATGAA 1024
 QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuG1 298
 DB 1025 TTGCACCTTAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACTGGG 1084
 QY 298 yPheSerLysValPheGlnAlaGluAenProPheAspPheMetGluAenIleSerLeuG1 318
 DB 1085 TTTTAGCAAGGTTTTCAGAGTAGAAGAACCCATTGACTTTATGGAGAAATATTTCACTGGA 1144
 QY 318 uGlyLysThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
 DB 1145 AGGAAGACTAATCTTTCAGAGAGAGTAGGCGAGTATCAGAGGATGGAGTGATGTC 1204
 QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1205 AAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1244

RESULT 9 ADK70302

ID ADK70302 standard; cDNA; 2215 BP.

XX AC ADK70302;

XX DT 06-MAY-2004 (first entry)

XX DE Respiratory disease differentially expressed cDNA #38.

XX KW de; gene; cytostatic; respiratory; antiasthmatic; gene therapy;
 KW differential gene expression; respiratory disorder; lung cancer;
 KW chronic obstructive pulmonary disease; emphysema; asthma.

XX OS Homo sapiens.

XX XX WO2003101283-A2.

XX XX 11-DEC-2003.

XX XX 02-JUN-2003; 2003WO-US017409.

XX XX 04-JUN-2002; 2002US-0386005P.

XX XX (INCY-) INCYTE CORP.

XX XX Rickert PK, Krasnow R;

XX XX WPI; 2004-042945/04.

XX PT New combination comprising cDNAs and proteins that are differentially
 PT expressed in respiratory disorders, useful for diagnosing or treating
 PT respiratory diseases e.g. lung cancer, chronic obstructive pulmonary
 PT diseases or asthma.

XX PS Claim 1; SEQ ID NO 38; 343pp; English.

XX CC The invention relates to cDNA sequences that are differentially expressed
 CC in respiratory disorders or their complements or encoded proteins. The
 CC cDNAs and proteins are useful for diagnosing, treating or monitoring
 CC treatment of a subject with a respiratory disease including lung cancer,
 CC chronic obstructive pulmonary diseases, emphysema or asthma. The protein
 CC is also useful for screening molecules or compounds to identify at least
 CC one ligand which specifically binds the protein. It is also useful for
 CC preparing and purifying a polyclonal or monoclonal antibody. This

CC sequence corresponds to a cDNA of the invention.
 XX Sequence 2215 BP; 593 A; 476 C; 523 G; 623 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,42e-167 Length: 2215
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 12 Gaps: 1

US-10-698-228-1 (1-351) x ADK70302 (1-2215)

QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
 DB 188 CCTGAGCGGACCCGCGCTCTGGCCAGCAAGACCGCAGGAGGATCTTCAGGAGCCAC 247
 QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
 DB 248 GGAGCCGAAACTAAAGCAGCTGCCCGCGGTGGAGGATGAGCCGCTCTGAGAGAAA 307
 QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysG1 58
 DB 308 CCCC CGCGCTTTGTCACTCTCCCATCAGTACCATGATATCTGCAGATGTATAGAA 367
 QY 58 nAlaGlnAlaSerPheThrThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
 DB 368 GGCAGAGCTTCCTTTTGGACCGCGGAGGAGGTGGACCTCTCCAGGACATTTCAGCAGCTG 427
 QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
 DB 428 GGAATCCCTGAAACCCGAGGAGAGATATTATATATCCCATGTTCTTGGCTTTCTTCGAGC 487
 QY 98 aSerAspGlyIleValAsnGluValGluArgPheSerGlnGluValGlnValPr 118
 DB 488 AAGCGATGGCATAGTAAATGAAACTTGTGTGAGCGATTAGCCAGAAAGTTTCAGATTAC 547
 QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
 DB 548 AGAAGCCGCTGTTTCTATGCTTCAAAATGCCATGGAAACATACATCTTGAAATGTA 607
 QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
 DB 608 TAGTCTTCTTATTCACACTTACATAAAAGATCCCAAGAAAGGGAATTTCTCTCAATGC 667
 QY 158 aIleGluThrMetProTyrValLysLysIleAlaAspTrpAlaLeuArgTrpIleAlaAs 178
 DB 668 CATTGAAACGATGCCCTTGTGTCAAGAAAGCAGACTGGGCTTTCGCTGGATTGGGGA 727
 QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaValGluGluValPhePh 198
 DB 728 CAAGAGGCTACCTATGTGTGAACGTTGTAGCCCTTGTCTGAGTGAAGAGCAATTTCTT 787
 QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 DB 788 TTCGGTTCTTTTGGCTCGATATTCTGGCTCAAGAAACGAGGACTGATGCTGGCTCAC 847
 QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 DB 848 ATTTTCTTAATGAACCTTATTAGCAGAGATGAGGGTTTACACTGTGATTTTGTCTGCTGAT 907
 QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAl 258
 DB 908 GTTCAACACCTGGTACCAACCATCGAGAGAGAGTAAGAGAAATAATTATCAATGC 967
 QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 DB 968 TGTTCGGATAGAACAGGAGTTCCTCACTGAGGCCCTTGTCTGTGAGCTCATTTGGGATGAA 1027
 QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuG1 298
 DB 1028 TTGCACTCTAATGAAGCAATATGTTGTTGTGGCAGACAGACTTATGCTGGAACCTGGG 1087

QY 298 yPheSerLysValPheGlnAlaGluAenProPheAspPheMetGluAenLysSerLeuG1 318
 Db 1088 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCATTGACTTTTATGAGAGATATTTCACTGGA 1147
 QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
 Db 1148 AGGAAGAGCTTACTCTTTGAGAGAGAGTAGAGCGAGTATCAGAGAGTGGAGTGATGTC 1207
 QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1208 AAGTCCACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1247
 RESULT 10
 ADJ56536
 ID ADJ56536 standard; cDNA; 2216 BP.
 XX
 AC ADJ56536;
 XX
 DT 06-MAY-2004 (first entry)
 DE Murine cDNA differentially expressed in MYCN activated cells SeqID 342.
 XX
 KW mouse; murine; differential expression; transactivator; proto-oncogene;
 KW neuroblastoma; small cell lung cancer; cytostatic; gene therapy; ss;
 KW MYCN activated cell.
 XX
 OS Mus musculus.
 XX
 PN US2003119009-A1.
 XX
 PD 26-JUN-2003.
 XX
 PF 25-FEB-2002; 2002US-00084817.
 XX
 PR 23-FEB-2001; 2001US-0270784P.
 XX
 PA (STUA/) STUART S G.
 PA (NUCH/) NUCHTERN J G.
 PA (PLON/) PLON S E.
 PA (SHOH/) SHOHEIT J M.
 XX
 PI Stuart SG, Nuchtern JG, Plon SE, Shohet JM;
 XX WPI; 2003-635698/60.
 XX
 PT New genes regulated by MYCN activation, useful in gene therapy,
 PT particularly for treating a subject with e.g. neuroblastoma or other
 PT cancers, or for diagnosing, staging or monitoring the treatment of the
 PT cancer.
 XX
 PS Claim 1; SEQ ID NO 342; 27pp; English.
 XX
 CC This invention relates to novel isolated cDNAs that are differentially
 CC expressed in MYCN activated cells. Specifically, it refers to
 CC polynucleotide sequences that exhibit differential expression patterns in
 CC cells activated by the transactivator MYCN, where MYCN is a proto-
 CC oncogene that is amplified in neuroblastoma cells and is common in small
 CC cell lung cancers. The present invention describes these cDNA molecules
 CC as useful for in hybridisation assays to detect expression of nucleic
 CC acids (or complementary nucleic acids) in a present in a given sample, as
 CC well as for screening assays by identifying molecules or compounds that
 CC specifically bind the cDNA as a ligand and modulate function or activity.
 CC Accordingly, these compositions exhibit cytostatic activity and can also
 CC be used for gene therapy purposes. This polynucleotide sequence is a cDNA
 CC that is differentially expressed in MYCN activated cells, given in an
 CC exemplification of the invention. NOTE: This sequence does not appear in
 CC the printed specification but has been obtained in electronic format from
 CC the US Patent Office at
 CC ftp.seqdata.uspto.gov/sequence.html?DocID=20030119009.
 XX
 SQ Sequence 2216 BP; 593 A; 476 C; 524 G; 623 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.42e-167 Length: 2216
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 10 Gaps: 1
 US-10-698-228-1 (1-351) x ADJ56536 (1-2216)
 QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
 Db 189 CTTGAGCGGACCCCGCTCTGGCCAGACACCGGAGGAGGATCTTCCAGGAGCCAC 248
 QY 23 nGluSerGluLysSer-----AsnGluGluProLeuLeuArgLysSe 38
 Db 249 GGAGCCGAAACTAAAGCAGCTGCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAAA 308
 QY 38 rSerArgArgPheValPhePheProLlePhePheProLlePhePheMetTyrLysG1 58
 Db 309 CCCCAGCGCTTGTCTATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAGAA 368
 QY 58 nAlaGlnAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
 Db 369 GGCAGAGCTTCTTTTGGACCGCGAGAGGTGACCTCTCCAGAGCATTCAGACTG 428
 QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheLleSerHisLleLeuAlaPhePheAlaA1 98
 Db 429 GGAATCCCTGAACCCCGAGGAGAGATATTTATATATCCCATGCTTCTGGCTTTCTTTCGAGC 488
 QY 98 aSerAspGlyLleValAsnGluAenLeuValGluArgPheSerGlnGluValGlnValPr 118
 Db 489 AAGCGATGGCATAGTAAACTTGGTGGAGCGATTAGCCCAAGAAAGTTTCAGATTAC 548
 QY 118 oGluAlaArgCysPheTyrGlyPheGlnLleLeuLleGluAenValHisSerGluMetTy 138
 Db 549 AGAAGCCCGCTGTTCATGGCTTCCAAATTTGCCATGAAACATACATCTGAAATGTA 608
 QY 138 rSerLeuLeuLeuAspThrTyrLleArgAspProLysLysArgGluPheLeuPheAsnAl 158
 Db 609 TAGTCTTCTTATGACACTTACATAAAGATCCCAAGAAAGGAAATTTCTTCTCAATGC 668
 QY 158 alleGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
 Db 669 CATTGAAACGATGCTTGTGTCAAGAAAGAGCCAGACTGGCGCTTGCCTGATTTGGGGA 728
 QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePh 198
 Db 729 CAAAGAGGCTACCTATGGTGAACGTGTGTAGCTTTGCTGAGTGGAGGAGGATTTCTT 788
 QY 198 eSerGlySerPheAlaAlaLlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 Db 789 TTCGGTTCCTTTTGGCTGATATTTCTGGCTCAAGAAAGAGGACTGATGCTGGCTCAC 848
 QY 218 rPheSerAsnGluLeuLysSerArgAspGluGlyLysHisCysAspPheAlaCysLeuWe 238
 Db 849 ATTTCTTAATGAACCTTATTAGCAGATGAGGGTTTACACTGTGATTTCCTTGCCTGAT 908
 QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgLleLleValAspAl 258
 Db 909 GTTCAAAACCTGTGTGTAACCAACCATCGAGGAGAGAGTAAGAGAAATAATATCAATGC 968
 QY 258 aValLysLleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuLleGlyMetAs 278
 Db 969 TGTTCCGATGAGAACAGGAGTTCCTCACTGAGGCTTTCCTGCTGAGCTCATTTGGATGAA 1028
 QY 278 nCysIleLeuMetLysGlnTyrLleGluPheValAlaAspArgLeuLeuValGluLeuG1 298
 Db 1029 TTGCACCTAATGAAGCAATACATTGATTTGTGGCAGACAGACTTATGCTGGAATGGG 1088
 QY 298 yPheSerLysValPheGlnAlaGluAenProPheAspPheMetGluAenLleSerLeuG1 318
 Db 1089 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCATTGACTTTTATGAGAGATATTTCACTGGA 1148


```
QY 258 aVallylLeuGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db 1401 TGTTCGATAGAACAGAGTCTCTCACTGAGGCTTGGCTGTGAAGCTCAATGGGATGAA 1342
QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuG1 298
Db 1341 TTGCATCTAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAATGGG 1282
QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
Db 1281 TTTTAGCAAGGTTTTCAGAGTAGAGAACCAATTTGACTTTATGAGATATTTCACTGGA 1222
QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
Db 1221 AGGAAGACTTAATCTTTTGGAGAGAGAGTAGGCGAGTATCAGAGATGGAGTGATGTC 1162
QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1161 AAGTCCAACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1122

RESULT 12
ABL65414
ID ABL65414 standard; DNA; 2500 BP.
AC ABL65414;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:3751.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; Gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 26-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
```

```
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
XX (AVAL-) AVALON PHARM.
PA
XX Young PE, Augustus M, Carter KC; Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
XX Claim 1; SEQ ID NO 3751; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.7e-167 Length: 2500
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: Gaps: 1

US-10-698-228-1 (1-351) x ABL65414 (1-2500)
QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
Db 302 CCTGAGCGGAGCCCGGCTCTGCGCAGACGCGGAGGAGATCTTCAGGAGCCAC 361
QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 362 GGAGCCGAAACTAAGACAGCTGCCCGCGCGTGGAGGATGAGCGCTGCTGAGAGAAA 421
QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysG1 58
Db 422 CCCCCCGCGCTTTGTCTATCTTCCCATGAGTACCATGATATCTGGCAGATGTATAGAA 481
QY 58 nAlaGlnAlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
Db 482 GGCAGAGGCTTCTCTTTGGACCGCGGAGGAGGTGACCTCTCCAAAGACATTCAGCACTG 541
```

Qy 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaA 98
Db 542 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCCATGCTTTGCGCTTTTC 601
Qy 98 aSerAspGlyLeuValAsnGluLysValGluArgPheSerGlnGluValGlnValPr 118
Db 602 AAGCGATGGCAGATGTAATGAAACTTGTGGGCGGATTTAGCCAAAGATTCAGATTAC 661
Qy 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 662 AGAAGCCCGCTGTTCTATCGCTTCCAAATTTGCCATGCAAAACATACATTTCTGAATGTA 721
Qy 138 rSerLeuLeuLeuAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db 722 TAGTCTTCTTATGACACTTACATAAAAGATCCCAAGAAAGGAAATTTCTCTCAATGC 781
Qy 158 alIeGluThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrIleAlaAs 178
Db 782 CATTGAAACGATGCTTGTGTCAAGAAAGGCGAGACTGGCGCTTGGCTGGATGGGGA 841
Qy 178 pArgLysSerPheGlyGluArgValAlaAlaPheAlaValGluGlyValPhePh 198
Db 842 CAAGAGGCTACCTATGCTGAGAGCTGTGTAGCCCTTGTGCTGAGTGAAGGCATTTCTT 901
Qy 198 eSerGlySerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db 902 TTCGGTTCCTTTTGGTGGATATTTCTGGTCAAGAAAGGAGACTGTATGCTGCCTCAC 961
Qy 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db 962 ATTTTCTTAATGAATATTATGACAGATGAGGGTTTACACTGTGATTTTGTCTTGCCTGAT 1021
Qy 238 tPheGluTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db 1022 GTTCAACACCTGTGTACACAAACCATCGAGAGAGAGTAAAGAAATAATTCAATGC 1081
Qy 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db 1082 TGTTCGATAGACAGAGTTCTCTACTGAGGCTTGGCTGTGAAGCTCATTTGGATGAA 1141
Qy 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGl 298
Db 1142 TTGCACCTAATGAAGCAATACATTGATGTTGTGGCAGACAGACTTATGCTGGAACCTGG 1201
Qy 298 yPheSerIysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
Db 1202 TTTTAGCAAGGTTTTCAGGTAGAGAACCCATTTGACTTTATGAGAGATATTTCACTGGA 1261
Qy 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
Db 1262 AGGAAAGACTTAATCTTTTGAGAGAGAGTAGGCGAGTATCAGAGGATGGAGTGTATGTC 1321
Qy 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1322 AAGTCCAAACAGAGATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 13

ABL66517

ID ABL66517 standard; DNA; 2500 BP.

AC ABL66517;

XX ABL66517;

XX 15-MAY-2002 (first entry)

DT Lung cancer related gene sequence SEQ ID NO:4854.

DE Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;

XX Stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;

KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;

KW gene; ds.

XX Homo sapiens.

OS

XX

XX

XX

XX

XX

XX

XX

XX

PN WO200194629-A2.

XX 13-DEC-2001.

PD 30-MAY-2001; 2001WO-US010838.

XX 05-JUN-2000; 2000US-0209473P.

XX 05-JUN-2000; 2000US-0209531P.

XX 18-SEP-2000; 2000US-0233133P.

XX 18-SEP-2000; 2000US-0233617P.

XX 20-SEP-2000; 2000US-0234009P.

XX 20-SEP-2000; 2000US-0234034P.

XX 20-SEP-2000; 2000US-0234052P.

XX 22-SEP-2000; 2000US-0234509P.

XX 25-SEP-2000; 2000US-0234923P.

XX 25-SEP-2000; 2000US-0235077P.

XX 25-SEP-2000; 2000US-0235082P.

XX 25-SEP-2000; 2000US-0235134P.

XX 26-SEP-2000; 2000US-0235637P.

XX 26-SEP-2000; 2000US-0235638P.

XX 27-SEP-2000; 2000US-0235711P.

XX 27-SEP-2000; 2000US-0235720P.

XX 27-SEP-2000; 2000US-0235840P.

XX 27-SEP-2000; 2000US-0235863P.

XX 28-SEP-2000; 2000US-0236028P.

XX 28-SEP-2000; 2000US-0236032P.

XX 28-SEP-2000; 2000US-0236033P.

XX 28-SEP-2000; 2000US-0236034P.

XX 28-SEP-2000; 2000US-0236109P.

XX 28-SEP-2000; 2000US-0236111P.

XX 29-SEP-2000; 2000US-0236842P.

XX 29-SEP-2000; 2000US-0236891P.

XX 02-OCT-2000; 2000US-0237172P.

XX 02-OCT-2000; 2000US-0237173P.

XX 02-OCT-2000; 2000US-0237278P.

XX 02-OCT-2000; 2000US-0237294P.

XX 02-OCT-2000; 2000US-0237295P.

XX 02-OCT-2000; 2000US-0237316P.

XX 03-OCT-2000; 2000US-0237425P.

XX 03-OCT-2000; 2000US-0237598P.

XX 03-OCT-2000; 2000US-0237604P.

XX 03-OCT-2000; 2000US-0237606P.

XX 03-OCT-2000; 2000US-0237608P.

XX 01-NOV-2000; 2000US-0244867P.

XX 01-NOV-2000; 2000US-0245084P.

XX (AVAL-) AVALON PHARM.

XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;

XX Soppet DR, Weaver Z;

XX WPI; 2002-188264/24.

XX Screening for anti-neoplastic agent involves exposing cells to a chemical

XX agent to be tested for anti-neoplastic activity, and determining a change

XX in expression of a gene of a signature gene set.

XX Claim 1; SEQ ID NO 4854; 44pp; English.

XX The present invention describes a method (M1) for screening for an anti-

XX neoplastic agent. The method involves exposing cells to a chemical agent

XX to be tested for anti-neoplastic activity, determining a change in

XX expression of at least one gene (I) of a signature gene set, where (I)

XX comprises a sequence (S) selected from 8447 sequences (given in ABL61664

XX to ABL70110), or is at least 95% identical to (S), where a change in

XX expression is indicative of anti-neoplastic activity. (I) has cytostatic

XX activity and can be used in gene therapy. M1 can be used for screening an

XX anti-neoplastic agent, and can be used for producing a product which is

XX the data collected with respect to the anti-neoplastic agent as a result

XX of M1, and the data is sufficient to convey the chemical structure and/or

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC properties of the agent. M1 can be used in the treatment of cancer such
 CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
 CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
 CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
 CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
 CC tumour

XX Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;
 SQ

Alignment Scores:

Pred. No.: 1.7e-167 Length: 2500
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: Gaps: 1

US-10-698-228-1 (1-351) x ABL66517 (1-2500)

QY 4 ProGluArgProGluAlaAlaGlyLeuAspGluAspGlu-ArgSerSerSerSerAspThrAs 23
 DB 302 CTGAGCGGACCCGGGCTCTGCGCCAGCAGACCGGAGGAGGATCTTCCAGGAGCCAC 361
 QY 23 nGluSerGluLeuLeuSer-----AsnGluGluProLeuLeuArgLysSe 38
 DB 362 GGAGCGGAATAAGACGCTGCCCGCGCGGTGGAGGATGAGCCGCTGCTGAGAGAAA 421
 QY 38 rSerArgArgPheValIlePheProIleGlnTyProAspIleTrpLysMetTyLysGI 58
 DB 422 CCCCAGCGCTTGTCTATCTCCCATCGAGTACCATGATATCTGCGAGATGATAGAA 481
 QY 58 nLaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
 DB 482 GGCAGAGGCTCTCTTTGGACCGCCGAGGAGGTGACCTCTCAAGGACATTCAGCACTG 541
 QY 78 pAsnLysLeuLysAlaAspGluLysTyPheIleSerHisIleLeuAlaPhePheAla 98
 DB 542 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCATGTTCTGCGCTTCTTTCGAGC 601
 QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
 DB 602 AGCGATGGCATAGTAAGTAAAGAACTTGTGGAGCATTTAGCCAGAGAGTTCCAGATTAC 661
 QY 118 oGluAlaArgCysPheTyArgPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
 DB 662 AGAAGCCCGCTGTTTCTATGGCTTCCAAATTGCCATGCAAAACATACATTCTGAATGTA 721
 QY 138 rSerLeuLeuLeuAspThrTyTrileArgAspProLysLysArgGluPheLeuPheAsnAl 158
 DB 722 TAGTCTTCTTATTGACACTTACATAAAGATCCCAAGAAAGGGAATTTCTCTTCAATGC 781
 QY 158 aIleGluThrMetProTyTrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
 DB 782 CATTAAGACATGCTCTGTGTCAAGAGAGAGGAGAGCTGGCGCTTGGCTGGATTGGGA 841
 QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
 DB 842 CAAAGAGGCTACCTATGTTGAACGTGTGTAGCTTTTGGCTGCGAGTGGAGGCATTTTCTT 901
 QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 DB 902 TTCCGGTCTCTTTCGTCGATATTCGCTCAAGAAACGAGGACTGATGCGCTCGCCCTCAC 961
 QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 DB 962 ATTTCTAATGAATTTATAGCAGATGAGGGTTTACATGTTGATTTTCTGCTGCTGAT 1021
 QY 238 tPheGlnTyLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
 DB 1022 GTTCAACACCTGGTACCAACCATCGGAGGAGAGTAAAGAGAAATATATCAATGC 1081
 QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278

DB 1082 TGTTCGATAGAACAGAGGTTCTCTCACTGAGGCCTTGCCTGAGAGCTCATTTGGATGAA 1141
 QY 278 nCysIleLeuMetLysGlnTyTrileGluPheValAlaAspArgLeuLeuValGluLeuGI 298
 DB 1142 TTGCACCTAATGAAGCAATACATTGAGTTTGTGCGACAGACACTTATGCTGGAACGGG 1201
 QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGI 318
 DB 1202 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCCAATTGACTTTATGGAGAAATATTTCACTGA 1261
 QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyTrGlnArgPheAlaValMetAl 338
 DB 1262 AGGAAGACCTAATCTTTTGAGAAGAGAGTAGGCGAGTATCAGAGGATGGGAGTGATGTC 1321
 QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1322 AAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 14
 ABL65859
 ID ABL65859 standard; DNA; 2500 BP.
 XX
 AC ABL65859;
 XX
 DT 15-MAY-2002 (first entry)
 XX
 DE Lung cancer related gene sequence SEQ ID NO:4196.
 XX
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200194629-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US010838.
 XX
 PR 05-JUN-2000; 2000US-0209473P.
 PR 05-JUN-2000; 2000US-0209531P.
 PR 18-SEP-2000; 2000US-0233133P.
 PR 18-SEP-2000; 2000US-0233617P.
 PR 20-SEP-2000; 2000US-0234009P.
 PR 20-SEP-2000; 2000US-0234034P.
 PR 20-SEP-2000; 2000US-0234052P.
 PR 22-SEP-2000; 2000US-0234509P.
 PR 22-SEP-2000; 2000US-0234567P.
 PR 25-SEP-2000; 2000US-0234923P.
 PR 25-SEP-2000; 2000US-0234924P.
 PR 25-SEP-2000; 2000US-0235077P.
 PR 25-SEP-2000; 2000US-0235082P.
 PR 25-SEP-2000; 2000US-0235134P.
 PR 25-SEP-2000; 2000US-0235280P.
 PR 26-SEP-2000; 2000US-0235637P.
 PR 26-SEP-2000; 2000US-0235638P.
 PR 27-SEP-2000; 2000US-0235711P.
 PR 27-SEP-2000; 2000US-0235720P.
 PR 27-SEP-2000; 2000US-0235840P.
 PR 27-SEP-2000; 2000US-0235863P.
 PR 28-SEP-2000; 2000US-0236028P.
 PR 28-SEP-2000; 2000US-0236032P.
 PR 28-SEP-2000; 2000US-0236033P.
 PR 28-SEP-2000; 2000US-0236034P.
 PR 28-SEP-2000; 2000US-0236109P.
 PR 28-SEP-2000; 2000US-0236111P.
 PR 29-SEP-2000; 2000US-0236842P.
 PR 29-SEP-2000; 2000US-0236891P.
 PR 02-OCT-2000; 2000US-0237172P.
 PR 02-OCT-2000; 2000US-0237173P.
 PR 02-OCT-2000; 2000US-0237278P.

PR 02-OCT-2000; 2000US-0237294P.
 PR 02-OCT-2000; 2000US-0237295P.
 PR 02-OCT-2000; 2000US-0237316P.
 PR 03-OCT-2000; 2000US-0237425P.
 PR 03-OCT-2000; 2000US-0237598P.
 PR 03-OCT-2000; 2000US-0237604P.
 PR 03-OCT-2000; 2000US-0237606P.
 PR 03-OCT-2000; 2000US-0237608P.
 PR 01-NOV-2000; 2000US-0244867P.
 PR 01-NOV-2000; 2000US-0245084P.
 XX (AVAL-) AVALON PHARM.
 XX
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX WPI; 2002-188264/24.
 XX
 PT Screening for anti-neoplastic agent involves exposing cells to a chemical
 PT agent to be tested for anti-neoplastic activity, and determining a change
 PT in expression of a gene of a signature gene set.
 XX
 PS Claim 1; SEQ ID NO 4196; 44pp; English.
 XX
 CC The present invention describes a method (M1) for screening for an anti-
 CC neoplastic agent. The method involves exposing cells to a chemical agent
 CC to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening an
 CC anti-neoplastic agent, and can be used for producing a product which is
 CC the data collected with respect to the anti-neoplastic agent as a result
 CC of M1, and the data is sufficient to convey the chemical structure and/or
 CC properties of the agent. M1 can be used in the treatment of cancer such
 CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
 CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
 CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
 CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
 CC tumour
 XX
 SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.7e-167 Length: 2500
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 6 Gaps: 1
 US-10-698-228-1 (1-351) x ABL65859 (1-2500)
 QY 4 ProGluArgProGluAlaAlaGlyLeuAspGluArgSerSerSerSerAspThrAs 23
 DB 302 CTGTAGCGGGAGCCCGGCTCTGCGCAGCAGACCGCGGAGAGGATCTTCAGGAGCCAC 361
 QY 23 nGluSerGluIleLeuSer-----AsnGluGluProLeuLeuArgLysSe 38
 DB 362 GGAGCGGAAACTAAGCAGCTGCCCGCGGAGGAGTATGAGCCGCTGCTGAGAGAAA 421
 QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysG1 58
 DB 422 CCCCCCGCGCTTGTTCATCTCCCATCGATACCATGATATCTGCAGATGTATAGAA 491
 QY 58 nAlaGlnAlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
 DB 482 GCAGAGGGTCTCTTTTGACCGCGGAGGAGGTGTACCTCTCCAGACATTCAGCACTG 541
 QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
 DB 542 GGAATCCCTGAAACCGAGGAGAGATATTTATATATCCATGTTCTGCTTCTTTCAGC 601

QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnValGlnValPr 118
 DB 602 AAGCGATGGCATAGTAATAAAGCTTGTGTGAGCGATTAGCCAAAGATTTCAGATTAC 661
 QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
 DB 662 AGAAGCCCGCTGTTCATGCTTCCAAATTTGCCAATGCGAAGAACATACATTTCTGAATGA 721
 QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
 DB 722 TAGTCTTCTTATGACACTTACATAAAGATCCCAAGAAAGGGAATTTCTCTCAATGC 781
 QY 158 aileGluThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
 DB 782 CATTGAAACGATGCTTGTGTCAAGAAAGCGCAGACTGGCGCTTGGCTGGATTGGGA 841
 QY 178 pArgLysSerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePh 198
 DB 842 CAAAGAGGCTACCTTATGCTGAACGTGTGTAGCCTTTGCTCAGTGGAGGATTTCTT 901
 QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 DB 902 TTCCGGTCTTTTGGCTGATATTTCTGGCTCAAGAAACGAGACTGATGCTGGCTCAC 961
 QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 DB 962 ATTTTCTAATGAACCTTATTAGCAGATGAGGGTTTACACTGTGATTTGCTTGGCTGAT 1021
 QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
 DB 1022 GTTCAAAACCTGTGTACCAAAACCATCGAGGAGAGAGTAAGAGAAATAATTTATCAATGC 1081
 QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 DB 1082 TGTTCGATGAGACAGAGTTCCTCCTCAGTGGCGCTTGCCTGTAAGCTCATTTGGGATGAA 1141
 QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG1 298
 DB 1142 TTGCACCTTATGAGCAATACATTTAGTTTGTGCGCAGACAGACTTATGCTGGAAGTGG 1201
 QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
 DB 1202 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCATTTGACTTTATGGAGATATTTCACTGGA 1261
 QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
 DB 1262 AGGAAAGACTAACTTCTTTTGAGAAAGAGATAGGCGGAGTATCAGAGGATGGGAGTGTGTC 1321
 QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1322 AAGTCAACAGAGATTTCTTTTACCTTGGATGCTGACTTC 1361
 RESULT 15
 ABX10335
 ID ABX10335 standard; DNA; 2500 BP.
 XX
 AC ABX10335;
 XX
 DT 28-JAN-2003 (first entry)
 XX
 DE DNA encoding protein differentially regulated in prostate cancer #4.
 XX
 KW Prostate cancer; gene expression; differential regulation;
 KW molecular marker; drug target; cancer detection; cancer diagnosis;
 KW cancer staging; cancer grading; cancer assessing; cancer monitoring;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200281638-A2.
 XX
 PD 17-OCT-2002.

XX 08-APR-2002; 2002WO-US010824.
 XX 06-APR-2001; 2001US-0281731P.
 PR 06-APR-2001; 2001US-0281732P.
 XX (ORIG-) ORIGENE TECHNOLOGIES INC.
 PA
 XX
 PI Sun Z, Jay G;
 XX WPI: 2003-058520/05.
 DR P-PSDB; ABU07433.
 DR
 XX
 PT Novel genes which are differentially regulated in prostate cancer, useful
 PT for diagnosing prostate cancer in prostate tissue sample and assessing
 PT therapeutic or preventive intervention in prostate cancer patients.
 XX
 XX Claim 1; Page 71-72; 416pp; English.
 XX
 CC The invention describes genes (I) which are differentially regulated in
 CC prostate cancer. (I) is useful for diagnosing a prostate cancer in a
 CC sample comprising prostate tissue, which involves determining the number
 CC of target genes which are differentially-regulated in the sample, where
 CC the number is indicative of the probability that the sample comprises
 CC prostate cancer. (I) is useful for assessing a therapeutic or preventive
 CC intervention in a subject having a prostate cancer, which involves
 CC determining the expression levels in a sample comprising prostate tissue
 CC of target genes which are differentially-regulated in prostate cancer.
 CC Preferably, the expression levels of at least 10 genes are determined.
 CC (I) is also useful for identifying agents that modulate a biological
 CC activity of a polypeptide differentially-regulated in prostate cancer
 CC cells, which involves contacting a polypeptide differentially-regulated
 CC in prostate cancer cells with a test agent under conditions effective for
 CC the test agent to modulate a biological activity of the polypeptide, and
 CC determining whether the test agent modulates the biological activity. (I)
 CC is useful as molecular markers, as drug targets, and for detecting,
 CC diagnosing, staging, grading, assessing, monitoring, prognosticating,
 CC preventing or treating, determining predisposition to diseases and
 CC conditions especially relating to prostate cancer. (I) and its expression
 CC products are used in the diagnostic test to assay for presence of cancer
 CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
 CC blood etc. (I) is useful for assessing cancer e.g., to determine the type
 CC of cancer, its stage of development, the nature of genetic defect, etc.
 CC The polypeptide encoded by (I) can be used as target for therapy or drug
 CC discovery. (I) can also be used for expressing the polypeptide and thus
 CC for searching specific binding partners of the polypeptide. (I) is useful
 CC in therapeutic applications to treat prostate cancer. The identification
 CC of specific genes, and groups of genes, expressed in pathways
 CC physiologically relevant to prostate cancer permits the definition of
 CC functional and disease pathways and the delineation of targets in these
 CC pathways which are useful in diagnostic, therapeutic, and clinical
 CC applications. This sequence encodes a protein differentially regulated in
 CC prostate cancer
 XX
 SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1,7e-167 Length: 2500
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 8 Gaps: 1
 US-10-698-228-1 (1-351) x ABX10335 (1-2500)
 QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSer-SerSerAspThrAs 23
 DB 302 CTGAGCGGGAGCCCGGCTCTGCGCCAGCAAGACCGCGGAGGAGTCTTCAGGAGCCAC 361
 QY 23 nGluSerGluileLysSer-----AsnGluGluProLeuLeuArgLysSe 38
 DB 362 GGAGCGGAAAACTAAAGCAGCTGCCCGCGCGTGGAGGATGAGCGCGTCTGAGAGAAA 421

QY 38 rSerArgArgPheValIlePhePheProIleGlnTyPProAspIleTrpLysMetTyLysGl 58
 DB 422 CCCCCCGCGCTTGTCTATCTTCCCATCGAGTACCATCATCTCGCAGATGATATAAGAA 481
 QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
 DB 482 GCGAGAGGCTTCTCTTTGGACCGCGGAGGAGGTGACCTCTCCAAAGGACATTCAGCATG 541
 QY 78 pAsnLysLeuLysAlaAspGluLysTyPheIleSerHisIleLeuAlaPhePheAlaAl 98
 DB 542 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCCATGTTCTGGCTTCTTTTCGAGC 601
 QY 98 aSerAspGlyIleValAsnGluAenLeuValGluArgPheSerGlnGluValGlnValPr 118
 DB 602 AAGCGATGGCATAGTAATAAAGAACTTGGTGAGCGATTTAGCCCAAGAAAGTTTCAGATTAC 661
 QY 118 oGluAlaArgCysPheTyPheGlyPheGlnIleLeuIleGluAenValHisSerGluMetTy 138
 DB 662 AGAAGCCCGCTGTTCTATGCTTCCAAATTTGCCATGGAAACATACATTTCTGAATGTA 721
 QY 138 rSerLeuLeuLeuAspThrTyPheIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
 DB 722 TAGTCTTCTTATTCACACTTACATAAAGATCCCAAGAAAGGAAATTTCTTTCATGTC 781
 QY 158 aIleGluThrMetProTyPheValLysLysLysAlaAspTrpAlaLeuAenGlyTrpIleAlaAs 178
 DB 782 CATTTAAACGATGCTTGTGTCAAGAAAGAGGAGACTGGCGCTTGGCTGGATTTGGGGA 841
 QY 178 pArgLysSerThrPheGlyGluArgValAlaPheAlaAlaValGluGluValPhePh 198
 DB 842 CAAGAGGCTACCTTGGTGAACGTGTGTAGCTTTCCTCAGTGGAGGATTTCTTCTT 901
 QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 DB 902 TTCCGGTCTTTTGGCTGATTTCTGGCTCAAGAAACGAGACTGATGCTGGCTGCTAC 961
 QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 DB 962 ATTTCTTAATGAACCTTATAGCAGAGATGAGGGTTTACACTGTGATTTTGTGCTGAT 1021
 QY 238 tPheGlnTyLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
 DB 1022 GTTCAAAACACCTGGTACACAAACCATCGGAGGAGAGTAAAGAGAAATAATATCAATGC 1081
 QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 DB 1082 TGTTGGATAGAACAGGAGTCTCTACTGAGGCTTGGCTGTGAAGCTCATTTGGGATGAA 1141
 QY 278 nCysIleLeuMetLysGlnTyPheIleGluPheValAlaAspArgLeuLeuValGluLeuGl 298
 DB 1142 TTGCACCTTAATGAAGCAATACATTTGAGTTTGTGGCAGACAGACTTATGCTGGAACGG 1201
 QY 298 yPheSerLysValPheGlnAlaGluAenProPheAspPheMetGluAenIleSerLeuGl 318
 DB 1202 TTTTAGCAAGGTTTTCAGAGTAGAGAAACCCATTTGACTTTATGGAGATAATTTTCACTGA 1261
 QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyPheGlnArgPheAlaValMetAl 338
 DB 1262 AGGAAAGACTAACTTCTTTGAGAAAGAGATGAGGAGTATCAGAGGATGGGAGTATGATGTC 1321
 QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1322 AAGTCCAAACAGAGAAATTTTACCTTGGATGCTGACTTC 1361

Search completed: June 16, 2005, 11:40:40
 Job time : 660 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 16, 2005, 09:43:32 ; Search time 5171 Seconds
(without alignments)
3289.070 Million cell updates/sec.

Title: US-10-698-228-1

Perfect score: 1821
Sequence: 1 MGDPERPEAGLQDQERSSS.....QRFVNAETTDNVFTLDADF 351

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n.model -DRV=xlh
-Q=/cgn2_1/USPTO.spool/US10698228/runat_14062005_161416_28295/app.query.fasta_1.519
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bites -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228 @cgn_1_4200 @runat_14062005_161416_28295 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	100.0	1053	6	AR454868
2	1821	100.0	1053	6	AR454877
3	1821	100.0	1053	6	BD064764
4	1821	100.0	1053	6	BD064773

5	1821	100.0	1053	6	BD093077	BD093077	Novel pro
6	1821	100.0	1053	6	BD093086	BD093086	Novel pro
7	1821	100.0	1056	6	CQ714252	CQ714252	Sequence
8	1821	100.0	1056	9	AB166671	AB166671	Homo sapi
9	1821	100.0	1081	6	AR454870	AR454870	Sequence
10	1821	100.0	1081	6	BD064766	BD064766	Novel pro
11	1821	100.0	1081	6	BD093079	BD093079	Novel pro
12	1821	100.0	4955	6	AR454869	AR454869	Sequence
13	1821	100.0	4955	6	BD064765	BD064765	Novel pro
14	1821	100.0	4955	6	BD093078	BD093078	Novel pro
15	1821	100.0	4955	9	AB036063	AB036063	Homo sapi
16	1817	99.8	1601	6	BD156916	BD156916	Primer fo
17	1817	99.8	1601	6	AX877905	AX877905	Sequence
18	1817	99.8	1601	9	AK001965	AK001965	Homo sapi
19	1725	94.7	4532	10	BC058103	BC058103	Mus muscu
20	1627.5	89.4	171737	10	AC122379	AC122379	Mus muscu
21	1514	83.1	900	9	AB163437	AB163437	Homo sapi
22	1474	80.9	4571	9	BC042468	BC042468	Homo sapi
23	1460.5	80.2	1649	9	BC001886	BC001886	Homo sapi
24	1460.5	80.2	1653	9	BC030154	BC030154	Homo sapi
25	1460.5	80.2	1794	9	S40301	S40301	Homo sapien
26	1460.5	80.2	2500	6	AR300103	AR300103	Sequence
27	1460.5	80.2	2500	6	AR380825	AR380825	Sequence
28	1460.5	80.2	2500	6	AR38503	AR38503	Sequence
29	1460.5	80.2	2500	6	AR562011	AR562011	Sequence
30	1460.5	80.2	2500	6	AX333242	AX333242	Sequence
31	1460.5	80.2	2500	6	AX333687	AX333687	Sequence
32	1460.5	80.2	2500	6	AX334345	AX334345	Sequence
33	1460.5	80.2	2500	9	HSRR288	HSRR288	Sequence
34	1456.5	80.0	3039	9	AK092671	AK092671	Homo sapi
35	1445.5	79.4	3203	9	AK123010	AK123010	Homo sapi
36	1444	79.3	1538	5	CR762018	CR762018	Xenopus t
37	1444	79.3	1611	5	CR848103	CR848103	Xenopus t
38	1443	79.2	2111	10	MUSNRM2A	M14223	Mouse ribon
39	1439	79.0	1406	5	CR353642	CR353642	Gallus ga
40	1439	79.0	1647	5	BC080161	BC080161	Xenopus t
41	1438	79.0	217774	2	AC120603	AC120603	Rattus no
42	1431	78.6	1605	5	BC047975	BC047975	Xenopus 1
43	1429	78.5	1170	6	E14825	E14825	cdna encodi
44	1428.5	78.4	1600	5	BC041209	BC041209	Xenopus 1
45	1425.5	78.3	1328	5	DRU57965	U57965	Danio rerio

ALIGNMENTS

RESULT 1	AR454868	Sequence 2 from patent US 6682917.	1053 bp	DNA	linear	PAT 20-FEB-2004
LOCUS	AR454868	Sequence 2 from patent US 6682917.				
DEFINITION	AR454868	Sequence 2 from patent US 6682917.				
ACCESSION	AR454868	Sequence 2 from patent US 6682917.				
VERSION	AR454868.1	GI:42688823				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 1053)					
AUTHORS	Nakamura,Y., Arakawa,H. and Tanaka,H.					
TITLE	Protein Having a Ribonucleotide Reductase activity and a DNA thereof					
JOURNALS	Patent: US 6682917-A 2 27-JAN-2004;					
FEATURES	Location/Qualifiers					
source	1..1053					
	/organism="unknown"					
	/mol_type="genomic DNA"					

ORIGIN

Alignment Scores:	1.01e-170	Length:	1053
Pred. No.:	1821.00	Matches:	351
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	100.00%	Gaps:	0
DB:	6		

US-10-698-228-1 (1-351) x AR454868 (1-1053)

```
QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
  |||
  |||
  |||
Db 1 ATGGCGAGCCCGAAGAGCGGAGCGCGCGGCTGGATCAGATCAGATCATCTTCA 60
  |||
  |||
  |||
QY 21 AspThrAsnGluSerGluLeuLysSerAsnGluProLeuLeuArgLysSerSerArg 40
  |||
  |||
  |||
Db 61 GACACCAACGAAGTGAATAAAGTCAAAATAGAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
  |||
  |||
  |||
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
  |||
  |||
  |||
Db 121 CGTTTGTTCATCTTCCAAATCCAGTACCCTGATATTGGAAATGATATAACAGGCACAG 180
  |||
  |||
  |||
QY 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
  |||
  |||
  |||
Db 181 GCTTCTCTTGGACACAGAGAGGTCGACTTATCAAGAGGATCTCCCTCACTGGACCAAG 240
  |||
  |||
  |||
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
  |||
  |||
  |||
Db 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTTCAATCTTGGAGGCTTTTTCAGGCGAGTAT 300
  |||
  |||
  |||
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
  |||
  |||
  |||
Db 301 GGAATTTGTAATGAATAATTTGGGGAGCGCTTTAGTCAGAGGTCAGGTTCCAGAGGCT 360
  |||
  |||
  |||
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
  |||
  |||
  |||
Db 361 CGCTGTTTCTATGGCTTTCAATCTCATCGAGAATGTTCTCATCAGAGATGTACAGTTTG 420
  |||
  |||
  |||
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
  |||
  |||
  |||
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
  |||
  |||
  |||
QY 161 ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrPheIleAlaAspArgLys 180
  |||
  |||
  |||
Db 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCTTCGATGGATGGATAGAGATAAAA 540
  |||
  |||
  |||
QY 181 SerThrPheGlyGluArgValValAlaPheAlaValAlaGluValPhePheSerGly 200
  |||
  |||
  |||
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTAGAGAGAGTTTCTTCTCAGGA 600
  |||
  |||
  |||
QY 201 SerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
  |||
  |||
  |||
Db 601 TCTTTTGTCTGCTATATTCTGGCTAAAGAGAGAGGTCCTTATGCCAGGACTCACTTTTTC 660
  |||
  |||
  |||
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
  |||
  |||
  |||
Db 661 AATGAACCTCATCAGCAGATGAAGGACTTCACTGTGACTTTTGGCTTGCCTGATGTCCAA 720
  |||
  |||
  |||
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
  |||
  |||
  |||
Db 721 TACTTAGTAATAAGCTTTCAGAAAGAGGTCAGGAGATCAATGTTGATGCTGCAAA 780
  |||
  |||
  |||
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
  |||
  |||
  |||
Db 781 ATTGAGCAGAGTTTTAAACAGAAGCTTTGCCAGTGGCTCATTTGGAATGAATTGCATT 840
  |||
  |||
  |||
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
  |||
  |||
  |||
Db 841 TTGATGAACACAGTACATTGAGTTGTAGCTGACAGATTAATTCTTGGAACTTTGGATTCTCA 900
  |||
  |||
  |||
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
  |||
  |||
  |||
Db 901 AAGGTTTTTCAGCAGAAAATCCTTTTGATTTTATGGAAACAAATTTCTTTAGAAAGAAA 960
  |||
  |||
  |||
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
  |||
  |||
  |||
Db 961 ACAATTTTCTTGAGAAAAGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
  |||
  |||
  |||
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
  |||
  |||
  |||
```

Db 1021 ACAGATAACGCTCTTCACCTTGGATGCAGATTTT 1053

RESULT 2

```
AR454877
LOCUS AR454877 1053 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 12 from patent US 6682917.
ACCESSION AR454877
VERSION AR454877.1 GI:42688832
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1053)
AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE Protein having a ribonucleotide Reductase activity and a DNA
therof
JOURNAL Patent: US 6682917-A 12 27-JAN-2004;
FEATURES
    Location/Qualifiers
        1..1053
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN
```

Alignment Scores:

```
Pred. No.: 1.01e-170 Length: 1053
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
```

US-10-698-228-1 (1-351) x AR454877 (1-1053)

```
QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
  |||
  |||
  |||
Db 1 ATGGCGAGCCCGAAGAGCGGAGCGCGCGGCTGGATCAGATCAGATCATCTTCA 60
  |||
  |||
  |||
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
  |||
  |||
  |||
Db 61 GACACCAACGAAGTGAATAAAGTCAAAATAGAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
  |||
  |||
  |||
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
  |||
  |||
  |||
Db 121 CGTTTGTCTGCTATATTCTGGCTAAAGAGAGGTCCTTATGGAAATGATATAACAGGCACAG 180
  |||
  |||
  |||
QY 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
  |||
  |||
  |||
Db 181 GCTTCTCTTGGACACAGAGAGGTCGACTTATCAAGAGGATCTCCCTCACTGGAAACAG 240
  |||
  |||
  |||
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
  |||
  |||
  |||
Db 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTTCAATCTTGGAGGCTTTTTCAGGCGAGTAT 300
  |||
  |||
  |||
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
  |||
  |||
  |||
Db 301 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGGTTCCAGAGGCT 360
  |||
  |||
  |||
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
  |||
  |||
  |||
Db 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCAGAAATGTTCTCATCAGAGATGTACAGTTTG 420
  |||
  |||
  |||
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
  |||
  |||
  |||
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
  |||
  |||
  |||
QY 161 ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrPheIleAlaAspArgLys 180
  |||
  |||
  |||
Db 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCTTCGATGGATGGATAGAGATAAAA 540
  |||
  |||
  |||
QY 181 SerThrPheGlyGluArgValValAlaPheAlaValAlaGluValPhePheSerGly 200
  |||
  |||
  |||
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTAGAGAGAGTTTCTTCTCAGGA 600
  |||
  |||
  |||
```

QY 201 SerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 DB 601 TCTTTTGTCTATATCTGGCTAAAGAGAGAGGCTTATGCGCAGGACTCACTTTTTC 650
 QY 221 ArgGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 DB 661 ATTGAACCTCATCAGCAGAGATGAAGGACTTCCTGTGACTTTTGGCTTGCCTGATGTTCCAA 720
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260
 DB 721 TACTTAGTAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTAGTGTCTCAA 780
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 DB 781 ATTGACGAGGAGTTTAAACAGAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTCATT 840
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 DB 841 TTGATGAACAGTACATGAGTTGTAGCTGCAGAGATTACTTGTGGAACCTTGGATTCTCA 900
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 DB 901 AGGTTTTTCAGCAGAGAAATCCTTTTGATTTTATGGAACAACTTTCTTAGAAGGAAA 960
 QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 DB 961 ACAAAATTTCTTTCAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTTATGGCAGAAACC 1020
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1021 ACAGATAACGCTCTTCACCTTGGATGCAGATTTT 1053
 RESULT 3
 BD064764
 LOCUS Novel protein and DNA thereof.
 DEFINITION BD064764
 ACCESSION BD064764.1 GI:22610367
 VERSION JP 2001269184-A/1.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
 TITLE Novel protein and DNA thereof
 JOURNAL Patent: JP 2001269184-A 1 02-OCT-2001;
 YUSUKE NAKAMURA,TAKEDA CHEMICAL INDUSTRIES LTD
 COMMENT OS Homo sapiens (human)
 PN JP 2001269184-A/1
 FD 02-OCT-2001
 PF 27-JUN-2000 JP 2000192401
 PI YUSUKE NAKAMURA,HIROFUMI ARAKAWA,HIROSHI TANAKA PC
 C12N15/09,A61K31/711,A61K38/00,A61K38/43,A61K45/00,A61K48/00, PC
 A61P35/00,
 PC A61P43/00,C07K16/32,C07K16/40,C12N1/15,C12N1/19,C12N1/21 PC
 ,C12N5/10,C12N9/02,
 PC C12P21/02,C12P21/08,G01N33/15,G01N33/50,G01N33/53/(C12N15/09,
 PC C12R1:91),
 PC (C12N15/09,C12R1:19),(C12N9/02,C12R1:91),(C12P21/02,C12R1:91),
 PC (C12P21/08,C12R1:91),C12N15/00,A61K37/02,A61K37/48,C12N5/00,
 PC (C12N15/00,C12R1:91),(C12N15/00,C12R1:19)
 CC Novel protein and DNA thereof
 FH Key Location/Qualifiers
 FT source 1..1053
 /organism='Homo sapiens (human)'.
 /organism='Homo sapiens'
 /mol_type='genomic DNA'
 /db_xref='taxon:9606'
 FEATURES
 source
 ORIGIN

Alignment Scores:
 Pred. No.: 1,01e-170 Length: 1053
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-10-698-228-1 (1-351) x BD064764 (1-1053)
 QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
 DB 1 ATGGCGCACCCGGAAGGCGGCGGCTGGATCAGGATGAGAGATCACTTCTCA 60
 QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
 DB 61 GACACCAACGNAAGTGAATAAAGTCAATCAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
 QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
 DB 121 CGGTTTGTCTATCTTCCAATCCAGTACCTCATATTTGGAATGTAATAACAGGCACAG 180
 QY 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrPheAsnLys 80
 DB 181 GCTTCTCTTGGACAGAGAGGTCGACTTATCAAGAGATCTCCCTCAGTCGAAACAG 240
 QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
 DB 241 CTTAAGCAGATGAGAAGTACTTCTCATCTCACAATCTTAGCTTTTTCAGCCAGTGTAT 300
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 DB 301 GGAATTCATAATGAAATTTGGTGGAGCGCTTATGTCAGGAGGTGACAGTTCACAGAGGCT 360
 QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 DB 361 CGCTGTTTCTATGCTTTCAAAATCTCATCGAGAATGTTCTACTCAGAGATGTACAGTTTG 420
 QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
 QY 161 ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
 DB 481 ACCATGCCCTATGTTAGAAAAAGACAGATTGGGCTTGGATGATGATGATGATGATGATA 540
 QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 DB 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTCTGCTGTAGAGGAGGATTTCTTCTCAGGA 600
 QY 201 SerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 DB 601 TCTTTTCTGCTATATTTCTGGCTAAGAAAGAGAGGCTTATGCGCAGGACTCACTTTTTC 660
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 DB 661 ATGAACCTCATCAGCAGAGATGAAGGACTTCAGTGTGACTTTGCTGTGATGTTCCAA 720
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 DB 721 TACTTAGTAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTAGTGTGATGTTCCAA 780
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 DB 781 ATTGACGAGGAGTTTAAACAGAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTCATT 840
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 DB 841 TTGATGAACAGTACATGAGTTGTAGCTGCAGAGATTACTTGTGGAACCTTGGATTCTCA 900
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 DB 901 AAGGTTTTTCAGGCAGAAAAATCCTTTTGATTTTATGGAACAACTTTCTTTAGAGGAAA 960

JOURNAL Patent: WO 0100799-A 1 04-JAN-2001;
TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,
HIROSHI TANAKA

COMMENT OS Homo sapiens (human)
PN WO 0100799-A/1
PD 04-JAN-2001
PP 27-JUN-2000 WO 2000JP004189
PR 28-JUN-1999 JP 99P 181131.06-JUL-1999 JP 99P 192391 PR
21-JAN-2000 JP 00P 017770
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N9/04, C12N15/53, C12N1/19, C12N1/19, C12N1/21, C12N5/10 PC
C12P21/02, A61K38/44
PC A61K48/00, C07K16/40, G01N33/53, C12Q1/26
CC
FH Key Location/Qualifiers.

FEATURES
source 1..1053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1,01e-170 Length: 1053
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-698-228-1 (1-351) x BD093077 (1-1053)

QY 1 MetGluValPheProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSerSer 20
DB 1 ATGGGCGACCCGGAAGCGGCGGCTGGATCAGATCAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluLeuIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
DB 61 GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
DB 121 CGGTTTGTTCATCTTCCATCCAGTACCTGATATTTGGAAATGTATTAACAGGCACAG 180
QY 61 AlaSerPheTrpTrpAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
DB 181 GCTTCCTCTTGACAGCAGAGAGGTCGACTTATCAAGAGATCTCCCTCAGTGAACAAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAlaSerAsp 100
DB 241 CTTAAGCAGATGAGAGTACTTCTCTCATCTTAGCTTTTTCAGCCAGTGTAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 301 CGAATTGTAAATGAAATTTGGTGGAGCGCTTAGTCAGGAGTGCGAGGCTCCAGAGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTTCTATGGCTTTCAAAATTCATCGAATGTTTCACTCAGAGATGTACAGTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACACTTACATCAGATCCAGATCCAGAAAGGAAATTTTATTTATGCAATTGAA 480
QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
DB 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGCCTTGGATGGATAGCAGATAGAAA 540
QY 181 SerThrPheGluGluArgValIleAlaPheAlaAlaValGluGlyValPheSerGly 200
DB 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGTGCTGTAGAGAGAGTTCTTCTCAGGA 600
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220

Db 601 TCTTTTGTCTATATCTTGCTTAAAGAGAGAGTCTTATCCAGGACTCATTCTTTCC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACCTGTGACTTTTGTCTGATGTTCCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGCAAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 781 ATTGAGCAGGAGTTTAAACAGAGACCTTGCCAGTTGGCCTCATTTGGAATGAATTGCAT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGlyPheSer 300
DB 841 TTGATGAACAGTACATTTGAGTTGTAGCTACAGATTTACTTTGGAACTTGGATCTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 901 AAGGTTTTTCAGGCAGAAATCTTTTGATTTTATGGAACCATTTCTTTAGAGGAAAA 960
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 961 ACAAAATTTCTTTGAGAAACGAGTTTTCAGATATCAGCGTTTTCAGTATTATGCGAGAAC 1020
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1021 ACAGTAACGTCCTTCAACCTTGGATGCAGATTTT 1053

RESULT 6
BD093086 1053 bp DNA linear PAT 27-AUG-2002
LOCUS Novel protein and its DNA.
DEFINITION BD093086
ACCESSION BD093086
VERSION BD093086.1 GI:22638674
KEYWORDS WO 0100799-A/10.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1053)
AUTHORS Nakamura,Y., ARAKAWA,H. and TANAKA,H.
TITLE Novel protein and its DNA
JOURNAL Patent: WO 0100799-A 10 04-JAN-2001;
TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,
HIROSHI TANAKA
COMMENT OS Homo sapiens (human)
PN WO 0100799-A/10
PD 04-JAN-2001
PP 27-JUN-2000 WO 2000JP004189
PR 28-JUN-1999 JP 99P 181131.06-JUL-1999 JP 99P 192391 PR
21-JAN-2000 JP 00P 017770
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N9/04, C12N15/53, C12N1/19, C12N1/19, C12N1/21, C12N5/10 PC
C12P21/02, A61K38/44,
PC A61K48/00, C07K16/40, G01N33/53, C12Q1/26
CC

FEATURES
source 1..1053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1,01e-170 Length: 1053
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

DB: 6 Gaps: 0
US-10-698-228-1 (1-351) x BD093086 (1-1053)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 1 ATGGGGCGACCCGGAAAGCGCGAGAGCGCGCGGTGGATCAGGATCAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 61 GACACCAACGAAAGTGAATAAGTCAAATGAAGAGCCACTCTAAGAAAGAGTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysGlnAlaGln 60
DB 121 CGGTTTGTTCATCTTCCCAATCCAGTACCTGATATTTGGAATGTTAATAACAGGACAG 180
QY 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
DB 181 GCTTCTCTTGACAGCAGAGAGGTGACTTATCAAGAGTCTCTCTCAGTGAACAAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaLaserAsp 100
DB 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCAGTCTTGGAGCCAGTGAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATGAGAGTGTCTCAGAGATGACAGTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
QY 161 ThrMetProTyrValLysLysLysAlaAspThrAlaLeuArgTrpIleAlaAspArgLys 180
DB 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAA 540
QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
DB 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTGTAGAGAGGATTTCTTCTCAGGA 600
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 601 TCTTTTGTGCTATATTTCTGGCTTAAGAGAGAGGTCCTATGCGCAGACTCATTCTTC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 661 AATGAACCTCATCAGCAGAGATGAAGACTTCTACTGTGACTTTGCTTGCCTGATGTTCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB 721 TACTTAGTAAATTAAGCTTTCAAGAAAGAGGTCAGGAGATCATTTGTATGCTGTCAAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 781 ATGTAGCAGAGGTTTTTAACAGAGCTTCCAGTTGGCTCATTTGGAATGAATTCATTT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGTGGAACCTTGGATCTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 901 AAGGTTTTTTCAGGCGAAAAATCCTTTGATTTTATGGAATAACATTTCTTTAGAGGAAAA 960
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 961 ACAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTGTATGCGAGAAACC 1020
QY 341 ThrAspAsnValPheThrLeuAlaAspPhe 351

DB 1021 ACAGATAACGTCTTCCCTTGGATGCAGATTTT 1053

RESULT 7
CQ714252
LOCUS CQ714252 1056 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 186 from Patent WO02068579.
ACCESSION CQ714252
VERSION CQ714252.1 GI:42275109
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
JOURNAL Patent: WO 02068579-A 186 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
Location/Qualifiers
source 1..1056
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 1,01e-170 Length: 1056
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-698-228-1 (1-351) x CQ714252 (1-1056)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 1 ATGGGGCGACCCGGAAAGCGCGAGAGCGCGCGGTGGATCAGGATCAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
DB 61 GACACCAACGAAAGTGAATAAGTCAAATGAAGAGCCACTCTAAGAAAGAGTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysGlnAlaGln 60
DB 121 CGGTTTGTTCATCTTCCCAATCCAGTACCTGATATTTGGAATGTTAATAACAGGACAG 180
QY 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
DB 181 GCTTCTCTTGACAGCAGAGAGGTGACTTATCAAGAGTCTCTCTCAGTGAACAAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaLaserAsp 100
DB 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCAGTCTTGGAGCCAGTGAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATGAGAGTGTCTCAGAGATGACAGTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
QY 161 ThrMetProTyrValLysLysLysAlaAspThrAlaLeuArgTrpIleAlaAspArgLys 180
DB 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAA 540


```

Qy 181 SerThrPheGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 200
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTGTCTGCTAGAGGAGTCTTCTCTCAGGA 600

Qy 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGTCTCTATATTCTCGCTAAAGAGAGAGTCTTATGCCAGGACTCAGCTTTTTC 660

Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACCTTGTCTGCTGAGTTTCAA 720

Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluLeuIleValAspAlaValLys 260
Db 721 TACTAGTAATAAGCTTCAGAGAAAGGCTCAGGAGATCATTTGTTGCTGTCTCAA 780

Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 781 ATTGAGCAGGAGTCTTTTAAACAGAGGCTTGCAGTTGGCCTCATTTGGAATGAATTGCATT 840

Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 841 TTGATGAACAGTACATGAGTTGTTGAGTGTGACAGATTAATTTGAGAACTTTGGATTCTCA 900

Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 901 AAGTCTTTTTCAGGAGAGAAATCCTTTTGAATTTATGGAACAATTTCTTTAGAGGAAA 960

Qy 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 961 ACAATTTCTTTTTCAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020

Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1021 ACAGATAACGTCTTACCTGGTATGCAGATTTT 1053

RESULT 8
LOCUS AB166671
DEFINITION Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase
small subunit 2, complete cds.
ACCESSION AB166671
VERSION AB166671.1 GI:45259568
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Ugai,H. and Yokoyama,K.K.
REFERENCE
AUTHORS Ugai,H. and Yokoyama,K.K.
TITLE Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase
small subunit 2, complete cds
JOURNAL Published Only in Database (2004)
REFERENCE 2 (bases 1 to 1056)
AUTHORS Ugai,H. and Yokoyama,K.K.
TITLE Direct Submission
JOURNAL Submitted (05-MAR-2004) Hideyo Ugai, RIKEN BioResource Center, Gene
Engineering Division; 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074,
Japan (E-mail:ugai@rtc.riken.go.jp, Tel:81-29-836-3612,
Fax:81-29-836-9120)
FEATURES
source Location/Qualifiers
1..1056
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="p53R2_C19"
/cell_line="Saos-2"
/call_type="Osteogenic sarcoma"
1..1056
/gene="p53R2"
1..1056
/gene="p53R2"

```

```

/codon_start=1
/product="p53-inducible ribonucleotide reductase small
subunit 2"
/protein_id="BAD12267.1"
/db_xref="GI:45259569"
/translation="MGDPERPEAAGLDODERSSSDTNESEIKSNRPRLRKSSRRFVI
FPQIPDIWMYKQASQAFWTAEBVDLKDLPWNKLKADEKFIKSHILAFPAASDGI
VNEMLVERFSQEVQPEARCFYQQLIENHSEMYSLIDITYRDPKKREFLFAIE
TMPYVKKADWALRWIADKSTFGERVFAVEGFFSGFAAIKFKKRGMLPGIT
PSNELISDEGLHCDFACLFOYLVPKSEERREIIVDAVKIEQSFLEALPVGILG
MNCILMKQYIEFVADRLVLELGFVKVQFAENPFDFMENISLEGKTFFFEKRVSEYQRF
AVMAETDNDVFTLDAP"
ORIGIN
Alignment Scores: 1.01e-170 Length: 1056
Pred. No.: 1821.00 Matches: 351
Score: 1821.00
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-10-698-228-1 (1-351) x AB166671 (1-1056)
Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGluAspGluArgSerSerSer 20
Db 1 ATGGCGCAGCCGGAAGGCGGAGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Qy 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
Db 61 GACACCAACGAAAGTGAATAAAGTCAAATGAAGAGCCACTCCTAAGAAAGATGTTCTCGC 120
Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
Db 121 CGGTTTGTCTCTTTTCCAATCCAGTACCTGTATTTGGAAATGATATAACAGGCACAG 180
Qy 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db 181 GCTTCTCTCGACAGCAGAGAGTCTGACTTATCAAGGATCTCCCTCACTCGAACAAG 240
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTTACATCTTACCTTTTTCAGCCATTTTTCAGCAGCATGAT 300
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 301 GGAATTGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCAGAGGCT 360
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 420
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
Qy 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 481 ACCATGCCCTATGTTTAAAGAAAAAAGCAGATGCGGCTTCGCGATGGATAGCAGATAGAAAA 540
Qy 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGATAGAGAGTCTTCTTCTCAGGA 600
Qy 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGTCTATATTCTTGGCTAAAGAGAGAGGCTCTTATGCCAGGACTCAGCTTTTTC 660
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCTCATCTGTGACTTGTGCTTGCCTGATGTTCCAA 720
Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260

```



```

PI YUSUKE NAKAMURA,HIROFUMI ARAKAWA,HIROSHI TANAKA PC
C12N15/09,A61K31/711,A61K38/00,A61K38/43,A61K45/00,A61K48/00, PC
A61P35/00,
PC A61P43/00,C07K16/32,C07K16/40,C12N1/15,C12N1/19,C12N1/21 PC
,C12N5/10,C12N9/02,
PC C12P21/02,C12P21/08,G01N33/15,G01N33/50,G01N33/53//(C12N15/09,
C12R1/19),
PC (C12N15/09,C12R1/19),(C12N9/02,C12R1/19),(C12P21/02,C12R1/19),
PC (C12P21/08,C12R1/19),C12N15/00,A61K37/02,A61K37/48,C12N5/00,
PC (C12N15/00,C12R1/19),(C12N15/00,C12R1/19)
CC Novel protein and DNA thereof
FH Key Location/Qualifiers
FT source 1..1081 /organism="Homo sapiens (human)".

FEATURES
    source
        Location/Qualifiers
            1..1081
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1 04e-170 Length: 1081
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-698-228-1 (1-351) x BD064766 (1-1081)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
|
|
|
DB 20 ATGGGGCGACCGGAAGCGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA 79
|
|
|
QY 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
|
|
|
DB 80 GACCAACCAAGTAAGTAAAGTCAATGAAGAGCCACTCTCAAGAAAGAGTTCTCGC 139
|
|
|
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
|
|
|
DB 140 CGGTGTTGTCATCTTCCATCCAGTAGTACCTCGATATTGGAAATGATATAACAGGCACAG 199
|
|
|
QY 61 AlaserPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
|
|
|
DB 200 GCTTCCTCTGGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAACAG 259
|
|
|
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaIleAsp 100
|
|
|
DB 260 CTTTAAAGCAGATGAGAAGTACTTTCATCTCTACATCTTAGCCCTTTTTCAGCCAGTGAT 319
|
|
|
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
|
|
|
DB 320 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGAGGCTTCCAGAGCT 379
|
|
|
QY 121 ArgCysPheTrpGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
|
|
|
DB 380 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAAGTTCCTCAGATGATCAGATTG 439
|
|
|
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
|
|
|
DB 440 CTGATAGACACTTACATCAGATGCCAGAAAGGGAATTTTATTTATGCAATGAA 499
|
|
|
QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
|
|
|
DB 500 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTGGATGGATAGCAGATAGAAA 559
|
|
|
QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluValPhePheSerGly 200
|
|
|
DB 560 TCTACTTTTGGGGAAGAGTGGTGGCTTTGGCTGTGAGAGAGAGTTTCTTCTCAGGA 619
|
|
|
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
|
|
|

```

```

DB 620 TCTTTTGTCTATATTCTGGCTAAGAGAGAGGTCTTATGCCAGGACTCATCTTTTCC 679
|
|
|
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
|
|
|
DB 680 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 739
|
|
|
QY 241 TyrIleValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
|
|
|
DB 740 TACTTAGTAATAAGCCCTTCAGAGAAAGGGTCAGGGAGATCATTTGTTGATGCTGTCAA 799
|
|
|
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
|
|
|
DB 800 ATTGACGAGGAGTTTTHACAGAGCCCTTGCAGTTGGCCTCATTTGGAATGATGATTCAT 859
|
|
|
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGlyPheSer 300
|
|
|
DB 860 TTGATGAACACAGTACATTTGAGTTTGTAGCTGACAGATTTACTTTGTGAACTTGGATTCTCA 919
|
|
|
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGlyLys 320
|
|
|
DB 920 AAGGTTTTTTCAGGCAGAAAAATCCTTTTGATTTTATGGAAAAATTTCTTTTAGAAGGAAA 979
|
|
|
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
|
|
|
DB 980 ACAAATTTCTTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAAGAAC 1039
|
|
|
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
|
|
|
DB 1040 ACAGATAAGTCTTCACTTGGATGCGAGATTTT 1072

RESULT 11
BD093079 1081 bp DNA linear PAT 27-AUG-2002
LOCUS
DEFINITION Novel protein and its DNA.
ACCESSION BD093079
VERSION BD093079.1 Gi:22638667
KEYWORDS WO 0100799-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Nakamura,Y., Arakawa,H. and Tanaka,H.
Novel protein and its DNA
Patent: WO 0100799-A 3 04-JAN-2001;
TAKEDA CHEMICAL INDUSTRIES LTD,YOSUKE NAKAMURA,HIROFUMI ARAKAWA,
HIROSHI TANAKA
COMMENT
OS Homo sapiens (human)
PN WO 0100799-A/3
PD 04-JAN-2001
PF 27-JUN-2000 WO 2000JP004189
PR 28-JUN-1999 JP 99P 181131,06-JUL-1999 JP 99P 192391 PR
21-JAN-2000 JP 00P 017770
PI YUSUKE NAKAMURA,HIROFUMI ARAKAWA,HIROSHI TANAKA PC
C12N9/04,C12N15/53,C12N1/15,C12N1/19,C12N1/21,C12N5/10 PC
C12P21/02,A61K38/44,
PC A61K48/00,C07K16/40,G01N33/53,C12Q1/26
CC
FH Key Location/Qualifiers.
FEATURES
    source
        Location/Qualifiers
            1..1081
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1 04e-170 Length: 1081
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

```


QY 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 Db 845 TCTTTTCTGCTATATCTCGCTAAAGAGAGAGTCTTATGCCAGGACTCCTTTTTC 904
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db 905 AATGAACTCATCAGCAGATGAAGACTTCTACCTGTCATTTGCTGCTGATGTTCAA 964
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleLeuValAspAlaValLys 260
 Db 965 TACTTAGTAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTGATGCTGTCAA 1024
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 Db 1025 ATTGAGCAGGAGTTTAAACAGAGCCTTGCCAGTTGCCCTCATTTGCAATGAATTGCATT 1084
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 Db 1085 TTGATGAACAGTACATGAGTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 1144
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 Db 1145 AAGTTTTTCAGCAGAGAAATCCTTTGATTTTATGGAACATTTCTTTAGAGGAAA 1204
 QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 Db 1205 ACAAATTTCTTTGAGAAACAGATTTTCAGATATCAGCGTTTTCAGTTATGGCAGAAACC 1264
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1265 ACAGATAACGCTCTTACCTTGGATGCAGATTTT 1297

RESULT 13

LOCUS BD064765 4955 bp DNA linear PAT 27-AUG-2002
 DEFINITION Novel protein and DNA thereof.
 ACCESSION BD064765
 VERSION BD064765.1 GI:22610368
 KEYWORDS JP 2001269184-A/2.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 4955)
 Nakamura, Y., Arakawa, H. and Tanaka, H.
 Novel protein and DNA thereof
 Patent: JP 2001269184-A 2 02-OCT-2001;
 YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD
 OS Homo sapiens (human)
 PN JP 2001269184-A/2
 PD 02-OCT-2001
 PF 27-JUN-2000 JP 2000192401
 PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
 C12N15/09, A61K31/711, A61K38/00, A61K38/43, A61K45/00, A61K48/00, PC
 A61P35/00,
 PC A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC
 , C12N5/10, C12N9/02,
 PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53// (C12N15/09,
 PC C12R1:91),
 PC (C12N15/09, C12R1:19), (C12N9/02, C12R1:91), (C12P21/02, C12R1:91),
 PC (C12P21/08, C12R1:91), C12N15/00, A61K37/02, A61K37/48, C12N5/00,
 CC (C12N15/00, C12R1:91), (C12N15/00, C12R1:19)
 CC Novel protein and DNA thereof
 FH Key Location/Qualifiers
 FT source 1. 4955
 FT /organism='Homo sapiens (human)'.
 FT Location/Qualifiers
 1. 4955
 /organism='Homo sapiens'
 /mol_type='genomic DNA'
 /db_xref='taxon:9606'

FEATURES
source

ORIGIN

Alignment Scores:
 Pred. No.: 7,39e-170 Length: 4955
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-10-698-228-1 (1-351) x BD064765 (1-4955)
 QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
 Db 245 ATGGCGACCCCGAAAGGCGGAGCGCGGCTGGATCAGATGAGATGATCACTTCA 304
 QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
 Db 305 GACACCAAGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
 QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
 Db 365 CGGTTTGTCTATCTTCCAATCCAGTACCTGATATTTGGAAAATGTATTAACAGGCACAG 424
 QY 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
 Db 425 GCITTCCTTCTGGACAGCAGAGAGTGCACCTTATCAAGAGATCTCCCTCACTGGAACAAG 484
 QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
 Db 485 CTTAAAGCAGATGAGAGTACTTCACTCTCACTCTAGCCTTTTTCAGCCAGTGTAT 544
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 Db 545 GGAATTTGTAATAAGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCAGAGGCT 604
 QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 Db 605 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAAGATGTTCTACTAGAGATGTACAGTTG 664
 QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPhePheAsnAlaIleGlu 160
 Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 724
 QY 161 ThrMetProTyrValLysLysLysAlaAspTyrPalaLeuArgTyrPalaLeuArgLys 180
 Db 725 ACCATGCCCTATGTTAAGAAAAAGCAGATGGGCTTGGGATGGATGATAGATAGAAA 784
 QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 Db 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTCTTCTTCTCAGGA 844
 QY 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 Db 845 TCTTTTGGCTGTATATCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 904
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db 905 AATGAACTCATCAGCAGATGAAGACTTCACTGTGACTTTGCTTGTGCTGATGTTCCAA 964
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 Db 965 TACTTAGTAATAAGCCTTCAAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 1024
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 Db 1025 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATTGCATT 1084
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 Db 1085 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 1144
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 Db 1145 AAGTTTTTCAGGAGAGAAATCCTTTGATTTTATGGAACATTTCTTTAGAGGAAA 1204

```

QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 1205 ACAAAATTTCTTTGAGAAAGCAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1264

QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1265 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1297

RESULT 14
BD093078 4955 bp DNA linear PAT 27-AUG-2002
LOCUS Novel protein and its DNA.
DEFINITION BD093078
ACCESSION BD093078
VERSION WO 0100799-A/2.
KEYWORDS Novel protein and its DNA
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4955)
AUTHORS Nakamura, Y., Arakawa, H. and Tanaka, H.
TITLE Novel protein and its DNA
JOURNAL Patent: WO 0100799-A 2 04-JAN-2001.
TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,
HIROSHI TANAKA
COMMENT OS Homo sapiens (human)
PN WO 0100799-A/2
PD 04-JAN-2001
PF 27-JUN-2000 WO 2000JP004189
PR 28-JUN-1999 JP 99P 181131, 06-JUL-1999 JP 99P 192391 PR
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC
.C12P21/02, A61K38/44.
PC A61K48/00, C07K16/40, G01N33/53, C12Q1/26
CC
FH Key Location/Qualifiers.
FEATURES
source
1. .4955
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred: No.: 7,39e-170 Length: 4955
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-698-228-1 (1-351) x BD093078 (1-4955)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 245 ATGGGGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 304

QY 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 305 GACACCAACGAAAGTGAATATAAGTCAATGAAGAGCCATCTCTAAGAAAGAGTTCTCGC 364

QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
DB 365 CGTTTGTTCATCTTCCAAATCCAGTACCTCCGTATTTGGAAATGATATAACAGGCACAG 424

QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
DB 425 GCTTCCTTCTGCACGACGAGAGGTCGACTATCAAGAGATCTCCTACTCTGACACAG 484

QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100

```

```

DB 485 CTTAAACGACATGAGAAGTACTTCTCATCTCTACATCTTAGCCCTTTTTCAGCCAGTGAT 544
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 545 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGAGTTCCAGAGGCT 604
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 605 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAATGTTTCTACTCAGAGATGTACAGTTG 664
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGNA 724
QY 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
DB 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATGGGCCCTTGCATGGATGACGATAGAAAA 784
QY 181 SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 200
DB 785 TCTACTTTTGGGAAAGAGTGGTGGCTTGTCTCTGTAGAAAGAGTTTCTTCTCAGGA 844
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 845 TCITTTGCTGCTATATTTCTGGCTAAGAGAGAGGTTCTATGCCAGGACTCATTCTTTCC 904
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 905 AATGAACCTCATCAGCAGATGAAGGACTTCACTGTGCTTGTGCTGCTGATGTTCCAA 964
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB 965 TACTTAGTAAATAGCCCTTCAGAGAAAGGTCAGGGAGAGTCAATTGTTGATGCTGTCAA 1024
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 1025 ATTGACGAGGAGTTTAAACAGAGCCCTTCCAGTGGCCCTCATTTGGAATGAATGCAIT 1084
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 1085 TTGATGAACAGTACATTGAGTTTGTAGTGCAGAGATTACTTGTGGAACCTTGGATTCTCA 1144
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 1145 AAGGTTTTCAGGCAGAAAAATCCTTTTGAITTTATGGAACAACATTCTTTTGAAGGAAA 1204
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 1205 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1264
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1265 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1297

RESULT 15
BD093078 4955 bp mRNA linear PRI 11-MAR-2000
LOCUS Homo sapiens p53R2 mRNA for ribonucleotide reductase, complete cds.
DEFINITION AB036063
ACCESSION AB036063
VERSION AB036063.1 GI:7229085
KEYWORDS p53R2; ribonucleotide reductase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Tanaka, H., Arakawa, H., Yamaguchi, T., Shiraishi, K., Fukuda, S.,
Matsumi, K., Takei, Y. and Nakamura, Y.
TITLE A ribonucleotide reductase gene involved in a p53-dependent
cell-cycle checkpoint for DNA damage
JOURNAL Nature 404 (6773), 42-49 (2000)
MEDLINE 20179179
PUBMED 10716435

```

```

REFERENCE 2 (bases 1 to 4955)
AUTHORS Tanaka,H., Arakawa,H. and Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (18-DEC-1999) Yusuke Nakamura, University of Tokyo,
Institute of Medical Science, Human Genome Center, Laboratory of
Molecular Medicine; 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639,
Japan (E-mail:yusuke@ims.u-tokyo.ac.jp, Tel:+81-3-5449-5372,
Fax:+81-3-5449-5433)

FEATURES             Location/Qualifiers
     source          1..4955
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8q23.1"
     gene            1..4955
                     /gene="p53R2"
     CDS              245..1300
                     /genes="p53R2"
                     /codon_start=1
                     /product="ribonucleotide reductase"
                     /protein_id="BAA92434.1"
                     /db_xref="GI:7229086"
                     /translation="MGDPPEPAAGLDQDERSSDTNESDKNEPFLLRKSSRFVFI
                     FPIQPDIMKMYKQASFTWAEVDLSKDLPHNKLKADEKVPISHLAPFAASDGI
                     VNENLVERFSQEVQVPKRCFTGYQILINVHSEMYSLIDTDIRPKRFLPNATIE
                     TWPYVKKADWALRWIADRKSTGERVAFPAAVEGFFSGSPAAI FWLKKRGLMPGIL
                     FNEELISRDEGLHGLDFACILMFOVLVKNKPSSEVRVRELIIVDAVKIEQFLTEALPVGLIG
                     MNCILMKQIETFEVADRLVLLVELGFSKVFQAEINPFDMEINISLEKTNFFPKRVSEYQRP
                     AVMAETIDNVITLUDAF"
     polyA_site       4933
                     /gene="p53R2"
                     /note="22 a nucleotides"

ORIGIN
Alignment Scores:
Pred. No.:          7,39e-170      Length:          4955
Score:              1821.00         Matches:         351
Percent Similarity: 100.00%         Conservative:    0
Best Local Similarity: 100.00%       Mismatches:     0
Query Match:        100.00%         Indels:         0
DB:                 9               Gaps:          0

US-10-698-228-1 (1-351) x AB036063 (1-4955)

QY      1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB      245 ATGGGCGACCCGGAAGCGCGGAAGCGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304

QY      21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
DB      305 GACACCAAGAGAGTGAATTAAGTCAATGAGAGCCACTCTTAGAAGAGAGTTCTCGC 364

QY      41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
DB      365 CGGTTTGTTCATCTTTCCATCCAGTACCCTCGATATTGGAAAATGATATAAACAGGCACAG 424

QY      61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
DB      425 GCTTCTCTTGGACAGCAGAGAGGTGCGACTTATCAAGAGATCTCCCTCCTCGAACCAAG 484

QY      81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
DB      485 CTTAAAGCAGATGAGAAGTACTTCATCTCTCACTCTTAGCCCTTTTTCAGCCAGTGAT 544

QY      101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB      545 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 604

QY      121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB      605 CGCTGTGTTCTATGGCTTTCAAATTCATCAGAAATGTTTCACCTCAGAGATGTACAGTTTG 664

```

```

QY      141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB      665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 724

QY      161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTyrIleAlaAspArgLys 180
DB      725 ACCATGCCCTATGTAGAAAAAAGCAGATGGGCTTGGGATGGATGATAGCAGATAGAAA 784

QY      181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
DB      785 TCTACTTTTGGGAAAGAGTGGTGGCTTGTGCTGTAGAGAGAGTTCCTTCTCTCAGGA 844

QY      201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB      845 TCTTTTGTGCTGCTATATTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 904

QY      221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB      905 ATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGTGCTGCTGATGTCCAA 964

QY      241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB      965 TACTTAGTAATAAGCCTTCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 1024

QY      261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB      1025 ATTGAGCAGGAGATTTTAAACAGAAAGCCTTGCAGTTGGCCTCATTTGGAATGAATTGCAT 1084

QY      281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB      1085 TTGATGAACAAGTAGTACATTTGAGTTGTAGCTGACAGATTACTTTGTGGAACCTTGGATTCTCA 1144

QY      301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB      1145 AAGTTTTTTCAGGCAGAAAATCCTTTTGATTTTATGGAACATTTCTTTAGAGGAAA 1204

QY      321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB      1205 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTGCAGTTTATGCGAGAAAC 1264

QY      341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB      1265 ACAGATAACGCTCTTCACCTTGGATGCGAGATTTT 1297

```

Search completed: June 16, 2005, 13:06:51
Job time : 5183 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 16, 2005, 17:22:59 ; Search time 213 Seconds
(without alignments)
2696.400 Million cell updates/sec

Title: US-10-698-228-1
Perfect score: 1821
Sequence: 1 MGDPERPEAAGLDQDERSS.....QRFVNAETTDNVFTLDADF 351

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DRV=xlh
-Q=/cgn2_1/USPOI.spool/US10698228/runat_14062005_161417_28327/app.query.fasta_1.519
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228@cgn_1_177@runat_14062005_161417_28327 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORE=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	100.0	1053	4	US-10-019-733-2
2	1821	100.0	1053	4	US-10-019-733-12
3	1821	100.0	1081	4	US-10-019-733-4
4	1821	100.0	4955	4	US-10-019-733-3
5	1460.5	80.2	2479	4	US-09-949-016-2025
6	1460.5	80.2	2500	4	US-09-962-665-9
7	1460.5	80.2	2500	4	US-09-923-655-1370
8	1460.5	80.2	2500	4	US-09-963-333-9
9	1460.5	80.2	2500	4	US-09-962-677-9
10	1460.5	80.2	2500	4	US-09-949-016-145
11	1210	66.4	14176	1	US-08-307-499-1
12	1210	66.4	14176	1	US-08-307-499-14

c

c	13	1210	66.4	14176	3	US-09-299-268-1	Sequence 1, Appli
	14	1210	66.4	14176	3	US-09-299-268-14	Sequence 14, Appl
	15	1131	62.1	1236	4	US-09-248-796A-3914	Sequence 3914, Ap
	16	1072.5	58.9	825	1	US-08-307-499-40	Sequence 40, Appl
	17	1072.5	58.9	825	3	US-09-299-268-40	Sequence 40, Appl
	18	984.5	54.1	1112	1	US-08-136-743B-1	Sequence 1, Appli
	19	969.5	53.2	1083	4	US-09-248-796A-3915	Sequence 3915, Ap
	20	768	42.2	481	3	US-08-905-223-125	Sequence 125, App
	21	674	37.0	656	3	US-09-328-111-791	Sequence 791, App
	22	599	32.9	449	4	US-09-270-767-14670	Sequence 14670, A
	23	558	30.6	419	3	US-08-905-223-129	Sequence 129, App
	24	543.5	29.8	11820	4	US-09-949-016-11887	Sequence 11887, A
	25	543.5	29.8	11826	4	US-09-949-016-13767	Sequence 13767, A
c	26	513.5	28.2	13706	4	US-09-902-540-1124	Sequence 1124, Ap
	27	509.5	28.0	1053	4	US-09-902-540-2666	Sequence 2666, Ap
	28	488.5	26.8	366	1	US-08-307-499-13	Sequence 13, Appl
	29	488.5	26.8	366	3	US-09-299-268-13	Sequence 4435, Ap
	30	322.5	21.6	297	4	US-09-313-294A-4435	Sequence 588, App
	31	381.5	21.0	1338	4	US-09-328-352-588	Sequence 3180, Ap
c	32	378.5	20.8	1707	4	US-09-252-991A-3180	Sequence 2787, Ap
	33	377.5	20.7	1431	4	US-09-252-991A-2787	Sequence 2692, Ap
	34	370	20.3	984	4	US-09-252-991A-2692	Sequence 319, App
	35	355	19.5	373	4	US-09-401-064-319	Sequence 337, App
c	36	355	19.5	373	4	US-09-401-064-337	Sequence 1, Appli
	37	338	18.6	3278	1	US-07-778-890A-1	Sequence 1, Appli
	38	335	18.4	1230025	4	US-09-198-452A-1	Sequence 76, Appl
	39	335	18.4	1230230	4	US-09-438-185A-1	Sequence 76, Appl
	40	330	18.1	124884	4	US-09-661-596A-76	Sequence 2, Appli
c	41	330	18.1	124884	4	US-09-913-514-1	Sequence 13, Appl
	42	330	18.1	125157	4	US-09-913-514-2	Sequence 2, Appli
c	43	317	17.4	3678	4	US-09-112-580-13	Sequence 13, Appl
	44	307	16.9	35100	2	US-08-770-379-19	Sequence 19, Appl
c	45	307	16.9	35100	3	US-08-757-669A-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-10-019-733-2
; Sequence 2, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-2

Alignment Scores:	6.06e-227	Length:	1053
Pred. No.:	1821.00	Matches:	351
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	100.00%	Gaps:	0
DB:	4		

US-10-698-228-1 (1-351) x US-10-019-733-2 (1-1053)

Oy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20


```
Db 1 ATGGGCGACCCGGAAGGCGGAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
Db 61 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
Db 121 CGGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAAAATGTATAAACAGGCACAG 180
QY 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
Db 181 GCTTCTCTTGGACAGCAGAAGAGGTCGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 241 CTTAAGCAGATCAGAAAGTACTTCACTCTCACATCTTAGCCTTTTGGAGCCAGTGAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 361 CGCTGTTCTTATGGCTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 420
QY 141 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 160
Db 421 CTGATAGACACTTACATCAGATCAGAGATCCCAAGAAAGGGAAATTTTATTAATGCATTTGAA 480
QY 161 ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrPheAlaAspArgLys 180
Db 481 ACCATGCCCTATGTTAAGAAAGAGCAGATTTGGGCTTGGATGATAGCAGATAGAGAAA 540
QY 181 SerThrPheGlyGluArgValAlaPheAlaValAlaValGluGlyValPhePheSerGly 200
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCGCTTTGCTGTGTAGAGAGATTTCTTCTCAGGA 600
QY 201 SerPheAlaAlaIlePheTyrLeuLysLysAlaAspTyrAlaLeuArgTyrPheAlaAspArgLys 220
Db 601 TCTTTTGGCTGCTATATCTGGCTAAAGAAAGAGAGGCTTATGCGAGAGCTCACCTTTTTC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACATCATCAGCAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 721 TACTTAGTAAATAAGGCTTCAGAAAGAAAGGTCAGGAGATCATCTGTATGCTGTCAAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 781 ATGAGCAGAGTGTATTAACAGAAAGGCTTGGCCAGTTGGGCTCATTTGGAATGAATGCATT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 841 TTGATGAACAGTACATTGATGTTGTAGCTGACAGATTACTTGTGGAATTTGATTTCTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnLysSerLeuGluGlyLys 320
Db 901 AAGGTTTTTTCAGCAGAAAAATCCTTTTGTATTTTATGMAAACATTTCTTTTAGAAGGAAA 960
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1021 ACAGATAACGTTCTTACCTTTGGATGCAGATTTT 1053
```

RESULT 2

US-10-019-733-12

```
; Sequence 12, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-12
```

```
Alignment Scores:
Pred. No.: 6,06e-227 Length: 1053
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
```

US-10-698-228-1 (1-351) x US-10-019-733-12 (1-1053)

```
QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 1 ATGGCGGACCCGGAAGGCGGAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
Db 61 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
Db 121 CGGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAAAATGTATAAACAGGCACAG 180
QY 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
Db 181 GCTTCTCTTGGACAGCAGAAGAGGTCGACTTATCAAGAGATCTCCCTCAGTGGAAACAAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 241 CTTAAGCAGATCAGAAAGTACTTCACTCTCACATCTTAGCCTTTTGGAGCCAGTGAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 420
QY 141 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 160
Db 421 CTGATAGACACTTACATCAGATCAGAGATCCCAAGAAAGGGAAATTTTATTAATGCATTTGAA 480
QY 161 ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrPheAlaAspArgLys 180
Db 481 ACCATGCCCTATGTTAAGAAAGAGCAGATTTGGGCTTGGATGATAGCAGATAGAGAAA 540
QY 181 SerThrPheGlyGluArgValAlaPheAlaValAlaValGluGlyValPhePheSerGly 200
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCGCTTTGCTGTGTAGAGAGATTTCTTCTCAGGA 600
```

Qy 201 SerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 Db 601 TCTTTTGGCTGCTATATCTGGCTAAAGAAAGAGAGGCTTATGCGAGGACTCATTCTTTTCC 660
 Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db 661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACCTTGTGCTTGCCTGATGTTCAA 720
 Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleLeuValAspAlaValLys 260
 Db 721 TACTTAGTAATAAGCTTCAGAAAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
 Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 Db 781 ATTGAGCAGGAGTCTTTTAAACAGAGCCTTGCCAGTTGGCCTCATTTGGAATGAAATGCAAT 840
 Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 Db 841 TTGATGAAACAGTACATTTAGTTGAGTTGTAGCTGACAGATTTCTTGTGGAACCTTGGATTCTCA 900
 Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 Db 901 AAGGTTTTTCAGCAGAGAAATCCTTTTGATTTTATGAAACATTTCTTTAGAGGAAA 960
 Qy 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 Db 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGGCTTTTGCAGTTATGGCAGAAACC 1020
 Qy 341 ThrAsnValPheThrLeuAspAlaAspPhe 351
 Db 1021 ACAGATAACGCTCTTCACCTTGGATGCAGATTTT 1053

RESULT 3

US-10-019-733-4
 ; Sequence 4, Application US/10019733
 ; Patent No. 6682917
 ; GENERAL INFORMATION:
 ; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
 ; TITLE OF INVENTION: New Protein and its DNA
 ; FILE REFERENCE: 2619WOOP
 ; CURRENT APPLICATION NUMBER: US/10/019, 733
 ; PRIOR FILING DATE: 2001-12-28
 ; PRIOR APPLICATION NUMBER: JP 11-181131
 ; PRIOR FILING DATE: 1999-06-28
 ; PRIOR APPLICATION NUMBER: JP 11-192391
 ; PRIOR FILING DATE: 1999-07-06
 ; PRIOR APPLICATION NUMBER: JP 2000-017770
 ; PRIOR FILING DATE: 2000-01-21
 ; NUMBER OF SEQ ID NOS: 14
 ; SEQ ID NO 4
 ; LENGTH: 1081
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION:
 US-10-019-733-4

Alignment Scores:
 Pred. No.: 6, 32e-227 Length: 1081
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-10-698-228-1 (1-351) x US-10-019-733-4 (1-1081)

Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
 Db 20 ATGGGGACCCGAAAGCGCGAAGCGCGGCTGGATCAGGATGAGATCATCTTCA 79
 Qy 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40

Db 80 GACACCAACGAAAGTCAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 139
 Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
 Db 140 CGGTTGTCTCTTCCAAATCCAGTACCTCATATTTGGAAAATGTATAAACAGGCACAG 199
 Qy 61 AlaserPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 Db 200 GCTTCCCTTCTCGACAGCAAGAGGTTGACTTATCAAGAGGATCTCCCTCACATGGAAACAG 259
 Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
 Db 260 CTTAAAGCAGATGAGAGTACTTCACTCTCATCTTAGCCTTTTTCAGCCAGTGTAT 319
 Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 Db 320 GGAATTGTAAATGAAAATTTGCTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 379
 Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 Db 380 CGCTGTCTTCTATGGCTTTCAAATTTCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 439
 Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 Db 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATGAA 499
 Qy 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 Db 500 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTTGGATGATGATAGCAGATAGAAA 559
 Qy 181 SerThrPheGlyGluArgValAlaPheAlaValGluGluValPhePheSerGly 200
 Db 560 TCTACTTTTGGGAAAGAGTGTGCTTGTCTGTAGAGGAGTTTCTTCTCAGGA 619
 Qy 201 SerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 Db 620 TCTTTTGTCTGCTATATCTTGGCTAAAGAAAGAGAGGCTTATGCCAGGACTCATTCTTTC 679
 Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db 680 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTGTCTGCTGATGTTCCAA 739
 Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 Db 740 TACTTAGTAATAAGCTTCAGAAAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 799
 Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 Db 800 ATTGAGCAGGAGTCTTTTAAACAGAAAGCCTTGCAGTTGGCCTCATTTGGAATGAAATGCAAT 859
 Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 Db 860 TTGATGAAACAGTACATTTGAGTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATTTCTCA 919
 Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGlyLys 320
 Db 920 AAGGTTTTTCAGGCAAGAAAATCCTTTTGATTTATGAAAACATTTCTTTAGAGGAAA 979
 Qy 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 Db 980 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGCTTTTTCAGAGTTATGGCAGAAACC 1039
 Qy 341 ThrAsnValPheThrLeuAspAlaAspPhe 351
 Db 1040 ACAGATAACGCTCTTCACCTTGGATGCAGATTTT 1072

RESULT 4

US-10-019-733-3
 ; Sequence 3, Application US/10019733
 ; Patent No. 6682917
 ; GENERAL INFORMATION:
 ; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
 ; TITLE OF INVENTION: New Protein and its DNA

FILE REFERENCE: 2619WOOP
CURRENT APPLICATION NUMBER: US/10/019,733
CURRENT FILING DATE: 2001-12-28
PRIOR APPLICATION NUMBER: JP 11-181131
PRIOR FILING DATE: 1999-06-28
PRIOR APPLICATION NUMBER: JP 11-192391
PRIOR FILING DATE: 1999-07-06
PRIOR APPLICATION NUMBER: JP 2000-017770
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 14
SEQ ID NO 3

LENGTH: 4955
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION:
US-10-019-733-3

Alignment Scores:
Pred. No.: 7,15e-226 Length: 4955
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-698-228-1 (1-351) x US-10-019-733-3 (1-4955)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 245 ATGGGGCGACCCGGGAAGGCGGAGCGCGCGCTGGATCAGATGATGATCTTCA 304
QY 21 AspThrAsnGlnSerGluLeuLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 305 GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTCTCGC 364
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
DB 365 CGGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAATGATATAACAGGCACAG 424
QY 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
DB 425 GCTTCTCTTGGACAGCAAGAGGTCGATTAACAAGATCTCCCTCACTGGAACAAG 484
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaLaserAsp 100
DB 485 CTTAAGCAGATGAGAAGTACTTCACTCTCACTTAGCCTTTTTCAGCCAGCGAT 544
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 545 GGAATTGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTCCAGAGGCT 604
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 605 CGCTGTGTTCTATGGCTTTCAATTTCACTGAGAATTTCACTCAGAGATGTACAGTTG 664
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 665 CTGATAGACACTTACATCAGATCCAGATCCAGAAAGGAAATTTTATTTATGCAATGAA 724
QY 161 ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
DB 725 ACCATGCCCTATGTTAAGAAAAAGAGATTTGGCCTTGGATGGATAGCAGATAGAAA 784
QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPheSerGly 200
DB 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAAGAGGTTTCTTCTCAGGA 844
QY 201 SerPheAlaAlaIlePheTyrLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 845 TCTTTTGTCTATATCTGGCTAAAGAGAGAGGCTTTATGCGAGAGCTCACTTTTTC 904
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240

DB 905 AATGAACATATCAGCAGAGATGAAGACTTCTACCTGTGACTTTGCTGCTGATGTTCCAA 964
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB 965 TACTTAGTAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 1024
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 1025 ATTGAGCAGGAGTGTAAACAGAGCCTTGCAGATTGCGCTCATTTGGAATGAATTCGAT 1084
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 1085 TTGATGAACACGATACATTGATGTTGTAGCTGACAGATTACTTTGGACACTTGGATCTCA 1144
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 1145 AAGGTTTTTCAGGACAGAAATCCTTTTGAATTTATGGAACAATTTCTTTAGAGGAAAA 1204
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 1205 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAGAAACC 1264
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1265 ACAGATAAACGTCTTCACTTGGATGCAGATTTT 1297

RESULT 5

US-09-949-016-2025
Sequence 2025, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: Fast-Seq for Windows Version 4.0
SEQ ID NO 2025
LENGTH: 2479
TYPE: DNA
ORGANISM: Human
US-09-949-016-2025

Alignment Scores:
Pred. No.: 2,06e-179 Length: 2479
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 4 Gaps: 1

US-10-698-228-1 (1-351) x US-09-949-016-2025 (1-2479)

QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerAspThrAs 23
DB 302 CCTGAGCGGACCCGGCTCTCGCCAGCAAGACCGGAGGAGGATCTTCCAGGAGCCAC 361
QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
DB 362 GGAGCCGAAACTAAAGCAGCTGCCCCCGCGTGGAGGATGAGCGCTGCTGAGAGAAA 421
QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysG 58
DB 422 CCCCCCGCGTTTGTCTATCTTCCCATCGAGTACCATGATATCTGCGAGATGTATAAGAA 481

```
QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 482 GGCAGAGCGCTCTCTTTGGACCGCGGAGAGGTGGACCTCTCCAGGACATTCAGCACTG 541
QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 542 GGAATCCCTGAAACCCGAGAGAGATATTTATATCCATGTTCTGGCTTTCTTTGAGC 601
QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 602 AAGCATGGCATAGTAATGAAACTTGTGTGGAGGATTTAGCCAGAGTTTCAGATTAC 661
QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 662 AGAAGCCCGCTGTTCTATGGCTTCCAAATGGCATGGAACATACATATTCGAAATGTA 721
QY 138 rSerLeuLeuLeuLeuAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 722 TAGTCTCTTTATTGACACTTACATAAAGATCCCAAGAAAGGGAATTTCTCTTCAATGC 781
QY 158 aIleGluThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTrpIleAlaAs 178
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 782 CATGGAACCATGCTTGTGTCAAGAAAGAGGAGACTGGGCTTGGCTGGATTTGGGA 841
QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePh 198
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 842 CAAAGAGCTACCTATGTTGACGTTGTGTAGCTTTGCTGCTGGAAGCATTTTCTT 901
QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 902 TTCGGTCTTTTGGCGTCGATATTCGCTCAAGAAACGAGGACTGATGCTGCGCTCAC 961
QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 962 ATTTTCTAATGAACTTATTAGCAGAGATGAGGTTTACACTGTGATTTTCTTGCCTGAT 1021
QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1022 GTTCAACACCTGGTACACAAACCTCGAGGAGAGTAGAAGAAATATTAATCAATGC 1081
QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1082 TGTTGGATAGAACAGAGTTCTCTCACTAGGCTTGGCTGTGAAGCTCATTTGGATGAA 1141
QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGl 298
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1142 TTGCACCTTAATGAAGCAATACATTTGATTTGTGGCAGACAGACTTATGCTGGAACCTGGG 1201
QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1202 TTTTAGCAAGTTTTCAGAGTAGAAGACCCATTGACTTTATGGAGAAATATTCACCTGGA 1261
QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1262 AGGAAGACTTAATCTTTTGAAGAGAGTAGGCGAGTATCAGAGGATGGAGTGTATGTC 1321
QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1322 AAGTCCCAACAGAGAAATCTTTTACCTTGGATGTGCTGCTTC 1361
```

RESULT 6

```
US-09-962-665-9
; Sequence 9, Application US/09962665
; Patent No. 6537759
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: POLYPOLYGLUTAMATE SYNTHETASE GENE SEQUENCE
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING THE
; TITLE OF INVENTION: TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015004
; CURRENT APPLICATION NUMBER: US/09/962,665
; CURRENT FILING DATE: 2001-09-24
```

```
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
```

```
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128..1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189..
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524..
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399..
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc_feature
; LOCATION: 1636..1738, 2259
; OTHER INFORMATION: n = c or t
US-09-962-665-9
```

```
Alignment Scores:
Pred. No.: 2,09e-179 Length: 2500
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 4 Gaps: 1
```

US-10-698-228-1 (1-351) x US-09-962-665-9 (1-2500)

```
QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlu-ArgSerSerSerAspThrAs 23
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 302 CCTGAGCGGAGACCGCGTCTGGCGCAGCAAGCCCGAGGAGGATCTTCCAGAGCCCCAC 361
QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 362 GGAGCCGAAACTAAAGCAGCTGCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAAA 421
QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGl 58
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 422 CCCCAGCGCGCTTGTCTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481
QY 58 nAlaGluAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 482 GGCAGAGCTTCTCTTTTGGACCGCGGAGGTTGACCTCTCNAAGGACATTCAGCAGCTG 541
QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 542 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCATGTTCTTGGCTTTCTTTGAGC 601
QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 602 AAGCATGGCATAGTAAATGAAACTTGTGTGGAGGATTTAGCCAGAGTTTCAGATTAC 661
QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 662 AGAAGCCCGCTGTTTCTATGCTTCCAAATTTGCCATGGAACATACATATTCGAAATGTA 721
QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

Db 722 TAGTCTTCTTATGACACTTACATAAAGATCCCAAGAAAGGAATTTCTTCAATGC 781
Qy 158 aileGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
Db 782 CATTTGAAACGATGCTTGTGTCAAGAAAGAGCAGACTGGCCCTTGGCTGGATGGGA 841
Qy 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaGluGlyValPhePh 198
Db 842 CAAAGAGGCTACCTATGTTGTAACGTTGTAGGCTTTGCTGCAAGGCAATTTCTT 901
Qy 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db 902 TTCGGTTCTTTTGGTGTGATATCTGGCTCAAGAAAGAGGACTGATGCTGGCTCAC 961
Qy 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db 962 ATTTCTTAATGAACCTTATTAGCAGAGATGAGGCTTTACACTGTGATTTTCTGCTGAT 1021
Qy 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db 1022 GTTCAACACCTGCTGACAAACCATCGGAGGAGAGTAAGAGAAATAATATCAATGC 1081
Qy 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db 1082 TGTTCCGATAGACAGAGTTCTCTACTGAGGCTTGCCTGTGAAGCTCATTTGGATGAA 1141
Qy 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAlaAspArgLeuValGluLeuG1 298
Db 1142 TTGCACTTAATGAACCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGAACCTGG 1201
Qy 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
Db 1202 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCATTTGACTTTATGGAGAAATATTTCACTGGA 1261
Qy 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
Db 1262 AGGAAAGACTAACTTTTGTAGAAGAGAGTAGGCGAGATCAGAGGATGGAGTGATGTC 1321
Qy 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1322 AAGTCCACACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 7

US-09-023-655-1370
; Sequence 1370, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREMITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:

CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1370:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2500 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: G36154
; US-09-023-655-1370

Alignment Scores:
Pred. No.: 2,09e-179 Length: 2500
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 4 Gaps: 1

US-10-698-228-1 (1-351) x US-09-023-655-1370 (1-2500)
Qy 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
Db 302 CCTGAGCGGACCGCGCTCTGCGCAGCAAGACCGCGAGGAGGATCTTCCAGGAGCCAC 361
Qy 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 362 GGAGCGGAACCTAAAGCAGCTGCCCCGGCGTGAGGATGAGCGCTGTGAGAGAAA 421
Qy 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysG1 58
Db 422 CCCCCGCGCTTGTCTATCTTCCCATCGATCATCATCATATCTGGCAGATGTATAAGAA 481
Qy 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
Db 482 GGCAGAGGCTTCTTTTGGACCGCGGAGGAGTTCCTCTCCAGGACATTCAGCAGCTG 541
Qy 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98
Db 542 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCATGTTCTGGCTTTCTTGGCAG 601
Qy 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 602 AAGCATGGCATAGTAATGAAACTTGGTGGAGCGATTTAGCCAAAGAGTTTCAGATTAC 661
Qy 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 662 AGAAGCCCGCTGTTTCTATGCTTCCAAATTTGCCAAACCAATACATCTTCTGAAATGA 721
Qy 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysArgGluPheLeuPheAsnAl 158
Db 722 TAGTCTTCTATTGACACTTACATAAAGATCCCAAGAAAGGGAATTTCTTCAATGC 781
Qy 158 aileGluThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
Db 782 CATTTGAAACGATGCTTGTGTCAAGAAAGAGCAGACTGGGCCCTTGGCTGGATGGGA 841
Qy 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
Db 842 CAAAGAGGCTACCTATGTTGTAACGTTGTAGGCTTTGCTGCAAGGCAATTTCTT 901
Qy 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db 902 TTCGGTTCTTTTGGTGTGATATTTCTGGCTCAAGAAAGAGGACTGATGCTGGCTCAC 961

Qy 218 rPheSerAsnGluLeuIleSerArgAspGluGluHisCysAspPheAlaCysLeuMe 238
Db 962 ATTTTCTTAATGAACCTTATTAGCAGAGATGAGGGTTTACACTGTGATTTTCTGCTCGCTGAT 1021
Qy 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluLeuIleValAspAl 258
Db 1022 GTTCAACACCTGGTACACAAACCATCGGAGGAGAGATGAGAGATAATATTATCAATGC 1081
Qy 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db 1082 TGTTGGATAGAACAGAGTTCTCACTAGAGGCTTGCCTGTGAAGCTCATTTGGATGA 1141
Qy 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGl 298
Db 1142 TTGCACCTTAATGAACATACATTTGAGTTGTGGCAGACAGACTTATGCTGGAACCTGGG 1201
Qy 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetCysAsnIleSerLeuGl 318
Db 1202 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCATTGTGACTTTATGGAGAAATATTTCACTGGA 1261
Qy 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
Db 1262 AGGAAGACTAATCTTTTGAGAGAGAGATGAGCGAGTATCAGAGGATGGAGATGATGTC 1321
Qy 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1322 AAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 8

US-09-963-333-9
; Sequence 9, Application US/09963333
; Patent No. 6664062
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE SYNTHASE GENE SEQUENCE VARIANCES
; TITLE OF INVENTION: HAVING UTILITY IN DETERMINING THE TREATMENT
; FILE REFERENCE: 11926-015002
; CURRENT APPLICATION NUMBER: US/09/963,333
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128, 1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc_feature
; LOCATION: 1636, 1738, 2259
; OTHER INFORMATION: n = c or t
; US-09-963-333-9

Alignment Scores: 2.09e-179 Length: 2500
Pred. No.: 1460.50 Matches: 277
Score: 1460.50
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: Gaps: 1
US-10-698-228-1 (1-351) x US-09-963-333-9 (1-2500)
Qy 4 ProGluArgProGluAlaAlaGlyLeuAspGlu-ArgSerSerSerAspThrAs 23
Db 302 CCTGAGCGGACCCCGCTCTGCGCAGCAAGACCGGAGGAGATCTTCAGAGAGCCAC 361
Qy 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 362 GGACCCGAAACTAAAGCAGCTGCCCCCGCGCTGGAGGATGAGCCGCTCTGAGAGAAA 421
Qy 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGl 58
Db 422 CCCC CGCGCTTTGTTCATCTTCCCATCGAGTACCATGATATCTGCAGATGTATAGAA 481
Qy 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
Db 482 GGCAGAGGCTTCTTTTGGACCGCGAGGAGGTTGACCTCTCNAAGGACATTCAGCAGCTG 541
Qy 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
Db 542 GGAATCCCTGAAACCCGAGGAGAGATATTATATCCCATGTCTGCGCTTCTTTTCAGC 601
Qy 98 aSerAspGlyIleValAsnGluValGluArgPheSerGlnGluValGlnValPr 118
Db 602 AAGCGATGCGCATAGTAAATGAAGAACTTGTGGAGCGATTTAGCCAGAAAGTTTCAGATTAC 661
Qy 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 662 AGAAGCGCGCTGTTCTATGCTTCCAAATGCGCATGCAAAACATACATCTCTGAAATGTA 721
Qy 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db 722 TAGTCTTCTTATTGACACTTACATAAAGATCCCAAGAAAGAGGAAATTTCTCTTCAATGC 781
Qy 158 aIleGluThrMetProTyrValLysLysAlaAlaAspTrpAlaLeuArgTrpIleAlaAs 178
Db 782 CATTTGAACGATGCTTGTGTGCAAGAGAGGACACTGGGCTTGGCGCTGGATTGGGGA 841
Qy 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
Db 842 CAAAGAGGCTACTATGCTGAACCGTGTGTAGCCCTTGTGCTGAGTGAAGGCAATTTCTT 901
Qy 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db 902 TTCCGGTTCTTTTGTGCTGATATTCTGCTCAAGAAACGAGGACTGATCCCTGGCCTCAC 961
Qy 218 rPheSerAsnGluLeuIleSerArgAspGluGluLeuHisCysAspPheAlaCysLeuMe 238
Db 962 ATTTTCTAATGAACCTTATTAGCAGAGATGAGGGTTTACACTGTGATTTTCTTGCCTGAT 1021
Qy 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db 1022 GTTCAACACCTGGTACACAAACCATCGGAGGAGAGATGAGAGAAATATTAATCAATGC 1081
Qy 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db 1082 TGTTGGATAGAACAGAGGATCTCTCACTGAGGCTTGGCTGTGCAAGCTCATTTGGGATGAA 1141
Qy 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGl 298
Db 1142 TTGCACCTTAATGAAGCAATACATTTGAGTTGTGGCAGACAGACTTATCTGGAACCTGGG 1201
Qy 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318

Db 1202 TTTTACGAGGTTTTCAGAGTAGAGAACCAATTTGACTTTATGGAGATATTTTCACTGGA 1261
 Qy 318 uGlyLysThrAsnPhenPheGluLysArgValSerGluTyrGlnArgPheAlaValMetal 338
 Db 1262 AGGAAAGACTAATCTTTTGAAGAAGAGTAGGCGAGTATCAGAGGATGGGAGTGATGTC 1321
 Qy 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1322 AAGTCCACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 9

US-09-962-677-9
 ; Sequence 9, Application US/09962677
 ; Patent No. 6759200
 ; GENERAL INFORMATION:
 ; APPLICANT: Stanton, Jr., Vincent P.
 ; TITLE OF INVENTION: THYMIDINE PHOSPHORYLASE GENE SEQUENCE
 ; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING
 ; TITLE OF INVENTION: THE TREATMENT OF DISEASE
 ; FILE REFERENCE: 11926-015003
 ; CURRENT APPLICATION NUMBER: US/09/962,677
 ; CURRENT FILING DATE: 2001-09-24
 ; PRIOR APPLICATION NUMBER: 09/658,659
 ; PRIOR FILING DATE: 2000-09-08
 ; PRIOR APPLICATION NUMBER: 09/596,033
 ; PRIOR FILING DATE: 2000-06-15
 ; PRIOR APPLICATION NUMBER: 09/357,743
 ; PRIOR FILING DATE: 1999-07-20
 ; PRIOR APPLICATION NUMBER: 09/357,024
 ; PRIOR FILING DATE: 1999-07-19
 ; PRIOR APPLICATION NUMBER: 60/093,484
 ; PRIOR FILING DATE: 1998-07-20
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 9
 ; LENGTH: 2500
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 128..1464
 ; OTHER INFORMATION: n = g or a
 ; NAME/KEY: misc feature
 ; LOCATION: 189
 ; OTHER INFORMATION: n = t or g
 ; NAME/KEY: misc feature
 ; LOCATION: 524
 ; OTHER INFORMATION: n = c or g
 ; NAME/KEY: misc feature
 ; LOCATION: 1399
 ; OTHER INFORMATION: n = t or a
 ; NAME/KEY: misc feature
 ; LOCATION: 1636..1738, 2259
 ; OTHER INFORMATION: n = c or t
 ; US-09-962-677-9

Alignment Scores:

Pred. No.:	2,09e-179	Length:	2500
Score:	1460.50	Matches:	277
Percent Similarity:	88.42%	Conservative:	36
Best Local Similarity:	78.25%	Mismatches:	35
Query Match:	80.20%	Indels:	6
DB:	4	Gaps:	1

US-10-698-228-1 (1-351) x US-09-962-677-9 (1-2500)

Qy 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
 Db 302 CTTGACGGGACCCGGGTCTTGGCCAGACGACCGGAGGAGGATCTTCCAGGAGCCAC 361
 Qy 23 nGluSerGluLeuLysSer-----AsnGluGluProLeuLeuArgLysSe 38
 Db 362 GGAGCCGAAAACTAAAGCAGCTGCCCGCGGTGGAGGATGAGCCGCTGCTGAGAGAAA 421

Qy 38 rSerArgArgPheValIlePhePheProfileGlnTyrProAspIleTrpLysMetTyrLysGl 58
 Db 422 CCCCCTGGCGTGTGTCTATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481
 Qy 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysSerLysAspLeuProHisTr 78
 Db 482 GGCAGAGGCTTCTTTTGGACCCGCGGAGGAGGTGACCTCTCNAAGGACATTCAGCACTG 541
 Qy 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
 Db 542 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCCATGTTCTGGCTTTCTTTGCGAGC 601
 Qy 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
 Db 602 AAGCGATGGCATAGTAATAAGAACTTGTGTGGAGCGATTTAGCCCAAGAAAGTTTCAGATTAC 661
 Qy 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
 Db 662 AGAAGCCCGCTGTTCTATGGCTTCCAAATTTGCCATGGAACATACATCTCGAATGTA 721
 Qy 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
 Db 722 TAGTCTTATTGACACTTACATAAAGATCCCAAGAAAGGAAATTTCTTCAATGC 781
 Qy 158 aIleGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
 Db 782 CATTGAAACGATGCTTGTGTCAAGAAAGAGGAGAGCTGGGCTTGGCTGGATGGGGA 841
 Qy 178 pArgLysSerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePh 198
 Db 842 CAAAGAGGCTACCTATGTTGAACGTGTTGTAGCTTTGTCAGTGGAGGCAATTTCTT 901
 Qy 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 Db 902 TTCGGTCTTTTGGTCGATATTTCTGCTCAAGAAACGAGGAGTATGCTTGGCTCAC 961
 Qy 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 Db 962 ATTTTCTAATGAACCTTATTAGCAGAGATGAGGGTTTACACTGTGATTTTGTCTGCTGAT 1021
 Qy 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
 Db 1022 GTTCAAAACACCTGGTACAAACCATCGAGGAGAGAGTAGAGAGAAATAATTAATCAATGC 1081
 Qy 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 Db 1082 TGTTGGATAGAACAGGAGTTCCTCACTGAGGCCCTTGGCTGTGAAGCTCATTTGGATGAA 1141
 Qy 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGl 298
 Db 1142 TTGCACCTTAATGAAGCAATFACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACCTGG 1201
 Qy 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
 Db 1202 TTTTAGCAGGTTTTCAGAGTAGAGAACCCATTTGACTTTATGGAGAAATATTTCACTGGA 1261
 Qy 318 uGlyLysThrAsnPhenPheGluLysArgValSerGluTyrGlnArgPheAlaValMetal 338
 Db 1262 AGGAAAGACTAATCTTTTGAAGAAGAGTAGGCGAGTATCAGAGGATGGGAGTGATGTC 1321
 Qy 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1322 AAGTCCACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 10

US-09-949-016-145
 ; Sequence 145, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF


```
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 145
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-145

Alignment Scores:
Pred. No.: 2,09e-179 Length: 2500
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 4 Gaps: 1

US-10-698-228-1 (1-351) x US-09-949-016-145 (1-2500)

Qy 4 ProGluArgProGluAlaAGlyLeuAspGluArgSerSerSerAspThrAs 23
Db 302 CCGAGCGGACCGCGCTCGTGGCCAGCAAGACCGCGAGGAGATCTTCCAGGAGCCAC 361
Qy 23 nGluSerGluLeuLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 362 GGAGCGGAACCTAAGCACTGCCCCCGCGTGGAGAGTAGGCGCTGCTGAGAGAAA 421
Qy 38 rSerArgArgPheValIlePheProLeuGlnTyrProAspIleTyrLysMetTyrLysG 58
Db 422 CCCCAGCGCTTTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481
Qy 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
Db 482 GCGAGAGGCTTCTTTTGGACCGCGAGAGGTTGACCTCTCCAAGGACATTCAGCACTG 541
Qy 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98
Db 542 GGAATCCCTGAACCCGAGAGAGATATTTATATCCAGTCTTGCTTCTTCTTGAGC 601
Qy 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 602 AAGCATGGCATAGTAATGAAACTTTGGTGGAGCGATTTAGCCAAGAGTTCCAGATTAC 661
Qy 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 662 AGAAGCCCGCTGTTCTATGGCTTCCAAATTTGCCATGCAAAACATACATTCGAAATGTA 721
Qy 138 rSerLeuLeuLeuAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db 722 TAGCTTCTTATTGACACTTACATAAAGATGCCAAAGAAAGGGAATTTCTTCAATGC 781
Qy 158 aIleGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
Db 782 CATTGAACCATGCTTGTCTCAAGAGAGGAGGAGCTGGGCTTGGCTGGATGGGA 841
Qy 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePh 198
Db 842 CAAAGAGGCTACCTGTCATCGTGTGTAGCTTTGCTGCTGCTGCAAGGCAATTTCTT 901
Qy 198 eSerGlySerPheAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db 902 TTCGGGTCTTTTGGCTGCTATTTCTGGCTCAAGAAACGAGGAGCTGATGCTGGCTCAC 961
Qy 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
```

```
Db 962 ATTTTCTAATGAACCTATTAGCAGAGATGAGGGTTTACACTGTGATTTTGTCTGCTGAT 1021
Qy 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db 1022 GTTCAAAACACCTGGTACACAAACCATCGAGGAGAGGTAAAGAGAAATAATTATCAATGC 1081
Qy 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db 1082 TGTTCCGATGAACACAGGAGTTCCTCCTGAGGCGCTTGGCTGTGAAGCTCATTTGGGATGAA 1141
Qy 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG 298
Db 1142 TTGCACCTTAATGAAGCAATACATTAAGTTGTGGCAGACACATTAATCTGGAACCTGGG 1201
Qy 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG 318
Db 1202 TTTTAGCAAGCTTTTTCAGAGTAGAGAACCCATTTGACCTTATGGAGAAATATTTCACCTGGA 1261
Qy 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
Db 1262 AGGAAAGACTTAATCTTTTGAGAGAGAGTAGGCGGAGTATCAGAGGATGGGAGTGATGTC 1321
Qy 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1322 AAGTCCACACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 11
US-08-307-499-1/c
; Sequence 1, Application US/08307499
; Patent No. 5651972
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,499
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
```

; LENGTH: 14176 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 3852..4226
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 4585..4887
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 5131..5310
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 5760..5912
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 6786..7130
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 10148..10513
 ; US-08-307-499-1

Alignment Scores:

Pred. No.: 1 37e-145 Length: 14176
 Score: 1210.00 Matches: 233
 Percent Similarity: 82.01% Conservative: 36
 Best Local Similarity: 71.04% Mismatches: 58
 Query Match: 66.45% Indels: 2
 DB: 1 Gaps: 1

US-10-698-228-1 (1-351) x US-08-307-499-1 (1-14176)

QY 24 GluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerSerArgArgPheVal 43
 DB 10842 AAAAATGATTTTATACAGAAAGTGGAGCCATATCTTCAAGAGTCTGATTCAGTTCGTT 10783
 QY 44 IlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGlnAlaSerPhe 63
 DB 10782 ATTTCCCTTAAGTATCATGATATCTGGAATGATGATTAACATCAGTGGCAGTTT 10723
 QY 64 TrpThrAlaGluGluValAlaPheSerLysAspLeuProHisTrpAsnLysLeuLysAla 83
 DB 10722 TGGACCGTTGAGCAAGTAGATTATCAAAAGATTTAGATGATGGGATAAATTAACATAA 10663
 QY 84 AspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAspGlyIleVal 103
 DB 10662 GACGAAAAATACTTTATAAACATATACTAGCATTTTTCATCTAGTGTGTTATGTA 10603
 QY 104 AsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAlaArgCysPhe 123
 DB 10602 AATGAGATTTACGGAAGATTTTATGTGTGATGACAGTGTTCAGAGCCAGCATGTTTC 10543
 QY 124 TyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeuLeuLeuAsp 143
 DB 10542 TATGATTTCAATAGCTATGGAATATTCATTCAGATGATGATATGATTTATTAATAGAT 10483
 QY 144 ThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGluThrMetPro 163
 DB 10482 ACATATGTAAGATAATAAGAAAAAATGCATTTATTAACGCTATAGAAACAATGGAA 10423
 QY 164 TyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLysSerThrPhe 183
 DB 10422 TCGTAAAAAAGAAAGCTGATTTGGGCGAGAAATGGATATCTAGCAACAAG--GTATAT 10366
 QY 184 GlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGlySerPheAla 203
 DB 10365 GGAAGAAGTAGTAGCATTTGAGCTGTGAGGGAATATCTTTCTGTTCAATTTGCT 10306
 QY 204 AlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeu 223

DB 10305 GCTATATTTTGGATAAAAAACAGAGATTGATCCCGGATTAAACATTTTCTAATGAACTA 10246
 QY 224 IleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGlnTyrLeuVal 243
 DB 10245 ATAAGTAGAGACGAAGGTTTACATTTGCGTGTGTTAATGTTTAAACATTTATTA 10186
 QY 244 AsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLysIleGluGln 263
 DB 10185 CATCCACCATCTAAGGAAGTTATACGTCGATAATCATGTGTCGGTTAATATAGAAAAG 10126
 QY 264 GluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIleLeuMetLys 283
 DB 10125 GAGTTTTGACAGTGTCTATTCGGTGGATCTTATAGGTATGAAATGTTGTTTAAATGCT 10066
 QY 284 GlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSerLysValPhe 303
 DB 10065 CAGTATATAGAAATTCGTCGAGATAGATTATTAACAGAGTTAGGTTGTGAAAAGT--CTC 10008
 QY 304 GlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLysThrAsnPhe 323
 DB 10007 AATGTATATATATCTTTTAGCTTTATGGAGTATATATCATAGAAAGGTAAAGCTAATTTT 9948
 QY 324 PheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThrThrAspAsn 343
 DB 9947 TTCGAACGACGAGTGTGATATATCAAAAGATGGGGTGTGTTTACAAATAAAGAAGAAAT 9888
 QY 344 ValPheThrLeuAspAlaAspPhe 351
 DB 9887 ATAATTTCTACGGATATAGATTTT 9864
 RESULT 12
 US-08-307-499-14
 ; Sequence 14, Application US/08307499
 ; Patent No. 5651972
 ; GENERAL INFORMATION:
 ; APPLICANT: Moyer, Richard W.
 ; APPLICANT: Vi uela, Eladio
 ; APPLICANT: Gibbs, E.P.J.
 ; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
 ; TITLE OF INVENTION: Live Vaccine Vector
 ; NUMBER OF SEQUENCES: 60
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: David R. Saliwanchik
 ; STREET: 2421 N.W. 41st Street, Suite A-1
 ; CITY: Gainesville
 ; STATE: Florida
 ; COUNTRY: U.S.A.
 ; ZIP: 32606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/307,499
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/908,241
 ; FILING DATE: 1-JUL-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/908,630
 ; FILING DATE: 29-JUN-1992
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/342,212
 ; FILING DATE: 21-APR-1992
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Saliwanchik, David R.
 ; REGISTRATION NUMBER: 31,794
 ; REFERENCE/DOCKET NUMBER: UP35.1.FWCCI
 ; TELECOMMUNICATION INFORMATION:


```

; Patent No. 6217882
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,268
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/901,127
; FILING DATE:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UP35.1.FWCCI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14176 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3852..4226
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4585..4887
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5131..5310
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5760..5912
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6786..7130
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 10148..10513
; US-09-299-268-1
; Alignment Scores:
; Pred. No.: 1.37e-145 Length: 14176

```

Db 9887 ATATTCTACGGATAGATTTT 9864
:::||||: ||| |||||
RESULT 14
US-09-299-268-14
; Sequence 14, Application US/09299268
; Patent No. 6217882
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,268
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/901,127
; FILING DATE:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14176 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 138..1460
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2456..2659
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2809..3030
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3070..3330
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3356..4180
; FEATURE:

; NAME/KEY: CDS
; LOCATION: 4392..5894
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6171..6398
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6447..6875
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6928..7431
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 7454..7858
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 7895..8155
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 8215..8682
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 8715..9539
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 9562..10272
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 10316..11908
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 11971..12780
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 12829..13107
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 13149..14171
US-09-299-268-14
Alignment Scores:
Pred. No.: 1.37e-145 Length: 14176
Score: 1210.00 Matches: 233
Percent Similarity: 82.01% Conservative: 36
Best Local Similarity: 71.04% Mismatches: 58
Query Match: 66.45% Indels: 2
DB: 3 Gaps: 1
US-10-698-228-1 (1-351) x US-09-299-268-14 (1-14176)
Qy 24 GluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArgArgPheVal 43
Db 3335 AAAAATGATTTTATACACGAATGGAGCTATTCTTCAAGAGCTGATTCTAGGTTTCGT 3394
Qy 44 IlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGlnAlaSerPhe 63
Db 3395 ATTTTCCTATTAAAGTATCATGATATCTGGAAAATGTATAACATCATCTGGCAAGTTT 3454
Qy 64 TrpThrAlaGluGluValAspLysSerLysAspLeuProHisTyrAsnLysLeuLysAla 83
Db 3455 TGGACCGTTGAAGAAAGTAGATTATCAAAAGATTTAGATGATTGGGATAAAATACTAAA 3514
Qy 84 AspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAlaSerAspGlyIleVal 103
Db 3515 GACGAAAATACTTTTATAAAAACATATACTAGCATTTTTCATCATCTAGTATGTTGA 3574
Qy 104 AsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAlaArgCysPhe 123
Db 3575 AATGAGATTTAGCGGAAAGATTTTATGTGATGATGATGATGATGATGATGATGATGAT 3634
Qy 124 TyrGlyPheGlnIleLeuGluAsnValHisSerGluMetTyrSerLeuLeuLeuAsp 143
Db 3635 TATGATTTTCAATAGCTATCGAAAATATTTCATTCAGAAATGATAGTTTATTATATAGAT 3694

QY 144 ThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGluThrMetPro 163
 DB 3695 ACATATGTAAAGATAATATAGAAAAAATGCAATTTATTTAAACGCTATAGAAACAATGGAA 3754
 QY 164 TyrValLysLysAlaAspTrpAlaLeuAUGTTrpIleAlaAspArgLysSerThrPhe 183
 DB 3755 TCGGTAAAGAAAGCTGATGGGCCAGAAATGATATCTAGCAACAG---GTATAT 3811
 QY 184 GlyGluArgValAlaPheAlaValAlaGluGluValPhePheSerGlySerPheAla 203
 DB 3812 CGAGAAAGAGTAGTACCAATTCAGCTGTGGAGGAATATCTTTCTGGTTCATTGGCT 3871
 QY 204 AlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeu 223
 DB 3872 GCTATATTTGGATAAAAAAGAGGATTCATGCCCGGATTAACATTTCTATGAACATA 3931
 QY 224 IleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGlnTyrLeuVal 243
 DB 3932 ATAAGTAGAGACGAAGGTTTACATTTGATTTTGCCTGTTTAAATGTTTAAACATTTATTA 3991
 QY 244 AsnLysProSerGluGluArgValArgGluIleIleValAlaValLysIleGluGln 263
 DB 3992 CATCCACCACTTAAGGAAGTTATAACGTCGATTAATCATTTGATGGGTTAATATAGAAAG 4051
 QY 264 GluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIleLeuMetLys 283
 DB 4052 GAGTTTTCACAGTTCCTATTCGGTGGATCTTATAGGTATGAATTTGTTTAAATGTCT 4111
 QY 284 GlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSerLysValPhe 303
 DB 4112 CAGTATATAGAAATTCGTCGACATAGATTATTAAACAGAGTTAGTTGTGAAAAGT--CTC 4169
 QY 304 GlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLysThrAsnPhe 323
 DB 4170 AATGTATATAATCCCTTTTAGCTTTATGGAGTATATATCTAGAGGTAAAGCTAATTTT 4229
 QY 324 PheGluArgValSerGluTyrGlnArgPheAlaValMetAlaGluThrThrAspAsn 343
 DB 4230 TTCCGAACGAGGTAGTCAATATCAAAAGATGGGGGTGTTTACAAATAAAGAGAGAAAT 4289
 QY 344 ValPheThrLeuAspAlaAspPhe 351
 DB 4290 ATATTTTCTACGGATATAGATTTT 4313
 RESULT 15
 US-09-248-796A-3914
 ; Sequence 3914, Application US/09248796A
 ; Patent No. 6747137
 ; GENERAL INFORMATION:
 ; APPLICANT: Keith Weinstock et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
 ; FILE OF INVENTION: 107196.132
 ; CURRENT FILING DATE: 1999-02-12
 ; CURRENT APPLICATION NUMBER: US/09/248,796A
 ; PRIOR FILING DATE: 1998-02-13
 ; PRIOR APPLICATION NUMBER: US 60/074,725
 ; PRIOR FILING DATE: 1998-02-13
 ; PRIOR APPLICATION NUMBER: US 60/096,409
 ; NUMBER OF SEQ ID NOS: 28208
 ; SEQ ID NO 3914
 ; LENGTH: 1236
 ; TYPE: DNA
 ; ORGANISM: Candida albicans
 US-09-248-796A-3914
 Alignment Scores:
 Pred. No.: 5,45e-137 Length: 1236
 Score: 1131.00 Matches: 210
 Percent Similarity: 78.21% Conservative: 52
 Best Local Similarity: 62.69% Mismatches: 67
 Query Match: 62.11% Indels: 6

DB: 4 Gaps: 3
 US-10-698-228-1 (1-351) x US-09-248-796A-3914 (1-1236)
 QY 23 AsnGluSerGluIleLysSerAsnGlu-----GluProLeuLeuArgLysSerSerArg 40
 DB 229 AACAGACACAAAGTTAAACCAAGAGAGTCCACGAGCAATTCCTTACACGACACAGACG 288
 QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
 DB 289 CGTCATACCATATATCGATAAAGTATCCGAGTATGGCAGTTCCTATAAAAGAGTCGTTA 348
 QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrp---Asn 79
 DB 349 GCCTCATCTTCGAGCTGCTGAAGAGCTTGAGCTGAGTAAAGATTGGATGATGGAATATAT 408
 QY 80 LysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAsaSer 99
 DB 409 AAGATGAACGAAACGAAAGATTCCTTATCTCTCGAGTTTGGCATCTTTTCTGTCATCG 468
 QY 100 AspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGlu 119
 DB 469 GATGGTATTCTTAATGAGAACTTGGTGGAAAAATTTCTGTGCAGAGTGCAAAATACCCGAG 528
 QY 120 AlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSer 139
 DB 529 GCTAAACTGCTTTACAGATTCAGATCATGATGAGAAACATTCATTTCAGAGACTTATTCG 588
 QY 140 LeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIle 159
 DB 589 TTGTTAATAGAAACATATTCAAGACCCCGAGACGACACTTTTGTGTTAAATGCTATT 648
 QY 160 GluThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTTrpIleAlaAspArg 179
 DB 649 GACATATCCCAATTATAAGAAAAAGCCGATTTGGGCAATCAGATGCATCCAGACGAG 708
 QY 180 LysSerThrPheGlyGluArgValAlaPheAlaValGluGluValPhePheSer 199
 DB 709 GATGCTTTGTATGCTGAGAGATTTGGTTCATTTGCTGCTGTTGAAGGTATATTTTTCAGT 768
 QY 200 GlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPhe 219
 DB 769 GGCTCGTTTGCAGCAATTTTCTGTTAAAGAAACGGGGTTTGATGCCGGGGTTGACATTT 828
 QY 220 SerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPhe 239
 DB 829 TCCATATGAATTAATTTGTCGAGACGAGGAGTTCACACAGATTATGCTGTTTATTTT 888
 QY 240 GlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaVal 259
 DB 889 TCATATTTAAAAAATAAACCTTCTTCAGAAAAATAATTGAGAAAAATCATCCGAGGCTGT 948
 QY 260 LysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCys 279
 DB 949 GATATCGAGAAAGAAATATTTTCAGCGATCGCTTCCAGTATCTCTTATTTGGGAATGAT 1008
 QY 280 IleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPhe 299
 DB 1009 GACTTGATGCGCCAGATGATTTGAATTTTGTGCTGATAGATTACTAGTAGCTTTTGGAAAT 1068
 QY 300 SerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGly 319
 DB 1069 AAAAAAGTACTACAAATGTTACCAACCCCATTTGATTTTATGAAAAATATTTCTCTGCTGT 1128
 QY 320 LysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGlu 339
 DB 1129 AAAACTAAATTTCTTTGAAAAAGAGAGTCTCTGATTATCAAAAGGCTGGTGTGATGAAAAA 1188
 QY 340 Thr-----ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1189 GTGAAAAATAAGAAATGAGAAAACTGGATTTGTTGATCAAGATTTTC 1233

Search completed: June 16, 2005, 18:40:38
Job time : 244 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 16, 2005, 13:06:58 ; Search time 703 Seconds
(without alignments)
3099.358 Million cell updates/sec

Title: US-10-698-228-1

Perfect score: 1821
Sequence: 1 MGDPRPEAAGLDQDRSSS.....QRFVMAETTDVNTLDADF 351

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6054689 seqs, 3103772919 residues

Total number of hits satisfying chosen parameters: 12109378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n_model -DRV=xlh
-Q=/cgn2_1/USPTO.spool/US10698228/runat_14062005_161418_28424/app.query.fasta_1.519
-DB=Published Applications NA -QWTF=fastap -SURFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPTXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10698228 -CGN 1 1 725 -runat_14062005_161418_28424
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10J_NEW_PUB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1821	100.0	1053	18	US-10-698-228-2	Sequence 2, Appli
2	1821	100.0	1053	18	US-10-698-228-12	Sequence 12, Appl
3	1821	100.0	1081	18	US-10-698-228-4	Sequence 4, Appli
4	1821	100.0	4955	17	US-10-172-118-71	Sequence 71, Appl
5	1821	100.0	4955	18	US-10-342-887-71	Sequence 71, Appl
6	1821	100.0	4955	18	US-10-698-228-3	Sequence 3, Appli
7	1460.5	80.2	1989	9	US-09-925-301-505	Sequence 505, App
8	1460.5	80.2	2216	15	US-10-084-817-342	Sequence 342, App
9	1460.5	80.2	2500	9	US-09-954-456-724	Sequence 724, App
10	1460.5	80.2	2500	9	US-09-954-456-1169	Sequence 1169, App
11	1460.5	80.2	2500	9	US-09-954-456-1827	Sequence 1827, App
12	1460.5	80.2	2500	18	US-10-641-643-1370	Sequence 1370, App
13	1460.5	80.2	2500	20	US-10-733-878-458	Sequence 458, App
14	1460.5	80.2	2500	21	US-10-843-641A-3751	Sequence 3751, App
15	1460.5	80.2	2500	21	US-10-843-641A-4196	Sequence 4196, App
16	1460.5	80.2	2500	21	US-10-843-641A-4834	Sequence 4834, App
17	1443	79.2	2113	21	US-10-764-420-1636	Sequence 1636, App
18	1425.5	78.3	1328	18	US-10-403-571-75	Sequence 75, Appl
19	1280	70.3	186854	21	US-10-872-156-34	Sequence 34, Appl
20	1230	67.5	1218	15	US-10-128-714-7245	Sequence 7245, App
21	1220.5	67.0	1292	17	US-10-369-493-27809	Sequence 27809, A
22	1192.5	65.5	1146	15	US-10-128-714-2245	Sequence 2245, App
23	1192	65.5	1242	16	US-10-032-585-6537	Sequence 6537, App
24	1176.5	64.6	963	17	US-10-369-493-34052	Sequence 34052, A
25	1168.5	64.2	1200	17	US-10-369-493-45857	Sequence 45857, A
26	1161.5	63.8	1173	17	US-10-369-493-46252	Sequence 46252, A
27	1161.5	63.8	1206	17	US-10-320-797-2145	Sequence 2145, App
28	1151	63.2	1314	15	US-10-128-714-1245	Sequence 1245, App
29	1151	63.2	1450	15	US-10-128-714-6245	Sequence 6245, App
30	1151	63.2	3314	15	US-10-128-714-245	Sequence 245, App
31	1151	63.2	3450	15	US-10-128-714-5245	Sequence 5245, App
32	1147.5	63.0	1586	18	US-10-424-599-73024	Sequence 73024, A
33	1133.5	62.2	1169	18	US-10-425-114-31361	Sequence 31361, A
34	1133.5	62.2	2992	20	US-10-425-115-31182	Sequence 31182, A
35	1132.5	62.2	1725	19	US-10-437-963-11109	Sequence 11109, A
36	1128.5	62.0	1255	19	US-10-767-701-13723	Sequence 13723, A
37	1125.5	61.8	1791	20	US-10-425-115-63321	Sequence 63321, A
38	1114.5	61.2	1248	18	US-10-424-599-126843	Sequence 126843, A
39	1113	61.1	1146	9	US-09-822-830A-174	Sequence 174, App
40	1097.5	60.3	1361	18	US-10-424-599-40471	Sequence 40471, A
41	1092.5	60.0	1569	20	US-10-425-115-15462	Sequence 15462, A
42	1062	58.3	957	20	US-10-653-047-43	Sequence 43, Appl
43	1043	57.3	1306	20	US-10-739-930-1378	Sequence 1378, App
44	1036.5	56.9	1420	17	US-10-320-797-1145	Sequence 1145, App
45	1036.5	56.9	3420	17	US-10-320-797-145	Sequence 145, App

ALIGNMENTS

RESULT 1

US-10-698-228-2
; Sequence 2, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06

```
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-2

Alignment Scores:
Pred. No.:          9.6e-217      Length:      1053
Score:              1821.00      Matches:     351
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%  Mismatches:  0
Query Match:        100.00%      Indels:      0
DB:                  18          Gaps:         0

US-10-698-228-1 (1-351) x US-10-698-228-2 (1-1053)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 1 ATGGGCGACCCGGGAAGCGCGAAGCGCGCGCTGGATCAGATGAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 61 GACACCAACCAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
DB 121 CGGTTTGTCTATCTTCCATCCAGTACCTCCCTGATATTGGAAATGATATAACAGGCACAG 180
QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
DB 181 GCTTCCTCTTGACACAGAGAGGTGCGACTTATCAAGGATCTCCTCACTGGAACAAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
DB 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCTACATCTTAGCCCTTTTTCAGCCAGTGAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTATTAATGCAATTGAA 480
QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
DB 481 ACCATGCCCTATGTTAAGAAAAAGAGATTTGGCCCTTGCATGATAGCAGATAGAAAA 540
QY 181 SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 200
DB 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGAGATTTTCTTCTCAGGA 600
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 601 TCTTTTGTCTATATTCTGGCTTAAGAGAGAGAGTCTTATGCGCAGAGCTACCTTTTTC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGTGCTGTATGCCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB 721 TACTTAGTAATAAAGCCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTTGATGTGTCAAA 780
```

```
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 781 ATTGAGCAGGAGTTTAAACAGAGAGCTTGCAGTTGGCCTCATTTGGAATGAATTCATT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATTGTGGAACCTTGGATCTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 901 AAGGTTTTTCAGGCAGAAAAATCCTTTTGAATTTATGGAAAAACATTTCTTTAAGAGGAAA 960
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 961 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTTGCAGTTTATGCGAGAAACC 1020
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1021 ACAGATTAACGTTCTTCACTTGGATGCGAGATTTT 1053

RESULT 2
US-10-698-228-12
; Sequence 12, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-12

Alignment Scores:
Pred. No.:          9.6e-217      Length:      1053
Score:              1821.00      Matches:     351
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%  Mismatches:  0
Query Match:        100.00%      Indels:      0
DB:                  18          Gaps:         0

US-10-698-228-1 (1-351) x US-10-698-228-12 (1-1053)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 1 ATGGGCGACCCGGGAAGCGCGAAGCGCGCGCTGGATCAGATGAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
DB 61 GACACCAACCAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
DB 121 CGGTTTGTCTATCTTCCATCCAGTACCTCCCTGATATTGGAAATGATATAACAGGCACAG 180
QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
DB 181 GCTTCCTCTTGACACAGAGAGGTGCGACTTATCAAGGATCTCCTCACTGGAACAAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
DB 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCTACATCTTAGCCCTTTTTCAGCCAGTGAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTATTAATGCAATTGAA 480
QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
DB 481 ACCATGCCCTATGTTAAGAAAAAGAGATTTGGCCCTTGCATGATAGCAGATAGAAAA 540
QY 181 SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 200
DB 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGAGATTTTCTTCTCAGGA 600
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 601 TCTTTTGTCTATATTCTGGCTTAAGAGAGAGAGTCTTATGCGCAGAGCTACCTTTTTC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGTGCTGTATGCCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB 721 TACTTAGTAATAAAGCCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTTGATGTGTCAAA 780
```

```
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleuAlaPhePheAlaIleSerAsp 100
DB 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTACATCTTAGCCCTTTTTCAGCCAGTGAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGAGGTTCCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTTTCTATGGCTTTCAAAATTCATCTCAGAGATGTTTCACCTCAGAGATGTACAGTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACACTTACATCAGATCCCAAGAAAGGAAATTTTATTTATGCAATTTGAA 480
QY 161 ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
DB 481 ACCATGCCCTATGTTAAGAAAGGAGATTTGGCCCTTGGATGGATAGCAGATAGAAAA 540
QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
DB 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGATTTCTTCTCAGGA 600
QY 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 661 AATGAACCTCATCAGCAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGTAGTTCCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB 721 TACTTAGTAAATAAGCCTTCAGAAAGAGAGGTGAGGATCAATGTTGATGCTGTCTCAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 781 ATTGACGAGAGTTTTAAACAGAGGCTTGCCAGTTGGCCTCATTTGGAATGAATGCATT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 841 TTGATGAAACAGTACATTCAGTTGTTAGCTGACAGATTACTTGTGAACTTGGATTTCTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 901 AAGGTTTTTTCAGCAGAAATCCTTTTGTATTTATGGAAACAACTTTCTTTAGAGAGGAAA 960
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 961 ACMAATTTCTTTGAGAAACAGTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1021 ACAGATAACGCTCTTCACCTTGGATGACAGATTTT 1053
```

RESULT 3

```
US-10-698-228-4
; Sequence 4, Application US/10698228
; Publication No. US2004007253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
```

```
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-4
Alignment Scores:
Pred. No.: 1e-216 Length: 1081
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0
```

US-10-698-228-1 (1-351) x US-10-698-228-4 (1-1081)

```
QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 20 ATGGCGCACCCGGAAGGCGGAGCGCGGCTGGATCAGGATGAGAGATCACTTCA 79
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
DB 80 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 139
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
DB 140 CGGTTTGTCTATCTTTTCCAATCCAGTACCTCATATTTCGAAATAATGATAAAACAGGCACAG 199
QY 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
DB 200 GCTTCCTCTTGGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCAGTGGACAAAG 259
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
DB 260 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCACATCTTTAGCCTTTTTCAGCCAGTGAT 319
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 320 GGAATTTGTAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGAGGTTCCAGAGGCT 379
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 380 CGCTGTTTCTATGGCTTTCAAAATTTCTATCAGAGATGTTCTCAGAGATGATACAGTTTG 439
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTGAA 499
QY 161 ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
DB 500 ACCATGCCCTATGTTAAGAAAGGAGTGGGCTTTCCTGCTAGAGAGGAGTTTCTTCTCAGGA 559
QY 181 SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 200
DB 560 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGAGTTTCTTCTCAGGA 619
QY 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 620 TCTTTTGTCTATATTTCTGGCTTAAAGAAAGAGAGGTCTTATGCGCAGGACTCACTTTTTC 679
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 680 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 739
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB 740 TACTTAGTAAATAAGCCTTTCAGAAAGAGGAGTCAATGTTGATGATCTGTCTGCAAA 799
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
```

```

Db      800 ATTGACGAGAGTTTTTAACAGAAAGCTTGCCAGTTGGCCCTCAATGGAAATGAATTGCATT 859
Qy      281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db      860 TTGATGAAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGGAACCTGGATTCTCA 919
Qy      301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db      920 AAGGTTTTTCAGCAGAGAAATCCTTTTGAATTTTATGGAACAACATTTCTTTAGAGGAAAA 979
Qy      321 ThrAsnPheGluLysArgValSerGluTyrGluArgPheAlaValMetAlaGluThr 340
Db      980 ACAAATTTCTTTGAGAAACAGATTTTACAGATATCAGCGTTTTCGAGTTATGGCAGAAACC 1039
Qy      341 ThrAsnValPheThrLeuAspAlaAspPhe 351
Db      1040 ACAGATAACGCTCTTCACCTTGGATGCAGATTTT 1072

RESULT 4
US-10-172-118-71
; Sequence 71, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Bernards, Rene
; APPLICANT: Van de Vijver, Marc
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; PRIOR FILING DATE: 2002-06-14
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AB036063
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-71

Alignment Scores:
Pred. No.: 1,1e-215 Length: 4955
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-10-698-228-1 (1-351) x US-10-172-118-71 (1-4955)
Qy      1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db      245 ATGGGGCAGCCCGGAAGCGCGGAAGCGCGGCTGGATCAGATGAGAGATCATCTTCA 304
Qy      21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerAtg 40
Db      305 GACACCAACGAAGTGAATTAAGTCAATGAAGACCACTCTCAAGAAAGATTTCTCGC 364
Qy      41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Db      365 CGTTTGTTCATCTTCCATCCAGTACCCTGATATTGGAAATGTATAAACAGGCACAG 424
Qy      61 AlaSerPheThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db      425 GCTTCCTCTGGACAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAAACAAG 484

```

```

Qy      81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaIaSerAsp 100
Db      485 CTTAAAGCAGATGAGAAAGTACTTCTCATCTCTCACATCTTTAGCCTTTTTCAGCCAGTGAT 544
Qy      101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db      545 GGAATGTTAAATGAAATTTTGTGAGCGCTTTTGTACAGGAGTGCAGGTTCCAGAGGCT 604
Qy      121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db      605 CGCTGTTTCTATGCTTTCAAATTTCTATCAGAAATGTTCACTCAGAGATGTACAGTTTG 664
Qy      141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db      665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAAATCAATTTGAA 724
Qy      161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTTrpIleAlaAspArgLys 180
Db      725 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTTGCCGATGATAGCAGATAGAAA 784
Qy      181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db      785 TCTACTTTTGGGAAAGAGTGGTGGCTTTCTGCTGTAGAAAGGAGTTTCTTCTCAGCA 844
Qy      201 SerPheAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db      845 TCTTTGCTGCTATATTCTGGCTAAAGAGAGAGGCTTATGCCAGGACTCACTTTTTC 904
Qy      221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db      905 AATGAATCTCATCAGCAGAGATGAAGGACTTTCACCTGTGACTTTGCTTGCTGTGATTTCAA 964
Qy      241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db      965 TACTTAGTAATAAGCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAA 1024
Qy      261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db      1025 ATTGAGCAGGAGTTTAAACAGAAAGCTTGCCAGTTGGCCTCATTTGGAATGAATTCATT 1084
Qy      281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db      1085 TTGATGAAACAGTACATTTGATGTTGTAGCTGACAGATTACTTGTGGAATCTGGATTTCTCA 1144
Qy      301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db      1145 AAGTTTTTTCAGGCAGAAATTCCTTTTGAATTTTATGGAACAATTTCTTTAGAGGAAAA 1204
Qy      321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db      1205 ACAATTTCTTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTATATGCGAGAAACC 1264
Qy      341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db      1265 ACAGATAACGCTCTTTCACCTTGGATGCAGATTTT 1297

```

RESULT 5

```

US-10-342-887-71
; Sequence 71, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887

```

```
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-71

Alignment Scores:
Pred. No.: 1,1e-215 Length: 4955
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-10-698-228-1 (1-351) x US-10-342-887-71 (1-4955)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 245 ATGGCGACCCCGAAAGCGGAAAGCGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
QY 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 305 GACACCAACGAAAGTAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 364
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
DB 365 CGTTTGTGCATCTTCCAAATCCAGTACCTCTCATCTCTACATCTTAGCCCTTTTTCGAGCGACAG 424
QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
DB 425 GCTTCCTCTTGACAGACGAGAGGTGCGATTCATCAAGAGATCTCCCTCACTGGAACAAG 484
QY 81 LeuLeuAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
DB 485 CTTAAAGCAGATGAGAAGTACTTCACTCTCATCTCTACATCTTAGCCCTTTTTCGAGCGAGTGAT 544
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCCAGAGGCT 604
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 605 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 664
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 724
QY 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
DB 725 ACCATCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGGCATGGATAGCAGATAGAAAA 784
QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
DB 785 TCTACTTTGGGGAAGAGTGGTGGCTTTGCTGTGTAGAGAGTTTCTTCTCAGGA 844
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 845 TCTTTTGTGCTATATCTTGCTTAAAGACAGAGGTCTTATGCCAGGACTCACTTTTTC 904
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 905 AATGAACATCATCAGCAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260

; TACTTAGTAATAAGCCCTTCAGAAAGGGTCAGGAGATCATTTGATGCTGCTCAA 1024
; IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
; ATTGAGCAGGAGTTTAAACAGAGCCCTTCCAGTTGGCTCATTCGAATGATTCATT 1084
; LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGlyPheSer 300
; TTGATGAAACACGATACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGAATCTCA 1144
; LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGlyLys 320
; AAGGTTTTTCAGGCGCAAAATCCTTTGATTTTATGGAACACATTTCTTTAGAAAGGAAA 1204
; ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
; ACAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1264
; ThrAspAsnValPheThrLeuAspAlaAspPhe 351
; ACAGATAACGTTCTTCACTTGCCTTGCATGACAGATTTT 1297

RESULT 6
US-10-698-228-3
; Sequence 3, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 3
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-3

Alignment Scores:
Pred. No.: 1,1e-215 Length: 4955
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-10-698-228-1 (1-351) x US-10-698-228-3 (1-4955)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 245 ATGGCGACCCCGAAAGCGGAAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
QY 21 AspThrAsnGluSerGluLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 305 GACACCAACGAAAGTAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 364
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
DB 365 CGTTTGTGCATCTTTCCAATCCAGTACCTCTGATATTTGAAAAATGTATAAACAGGACAG 424
QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
```

```
Db 425 GCTTCTCTCTGGACACAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAACAG 484
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaLaSerAsp 100
Db 485 CTTAAAGCAGATGAGAAGTACTTCTCTCATCTTACATCTTTTGGCAGCAGTGAT 544
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCCAGAGCT 604
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 605 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAATGTTCTCAGAGATGACAGTTG 664
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 665 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATGAA 724
Qy 161 ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTrpIleAlaAspArgLys 180
Db 725 ACCATCCCTATGTTAAGAAAAAGCAGATTTGGGCTTGGATCGATGATAGAGATAA 784
Qy 181 SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 200
Db 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAAGAGTTTCTTCTCAGGA 844
Qy 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 845 TCTTTTGTCTATATCTGGCTTAAGAGAGAGGCTTATGATCCAGGACTCATTTTTCC 904
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 905 AATGAACATCATCAGCAGAGATGAAGACITTCACGTGACTTTGCTTGCCTGATGCCAA 964
Qy 241 TyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAlaValLys 260
Db 965 TACTTAGTAAATAAGCTTTCAGAAAGAGGCTCAGGAGATCATTTGTGATGCTGTCAA 1024
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 1025 ATTGACGAGAGTTTAAACAGAGGCTTGCAGTTGGCTCATTTGGAATGATTCAT 1084
Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 1085 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTTCA 1144
Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 1145 AAGGTTTTTCAGCAGAGAAATCCTTTTGATTTTATGGAACAATTTCTTTAGAGGAAA 1204
Qy 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 1205 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1264
Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1265 ACAGATAACGTCTTCACTTGGATGACAGATTTT 1297
```

RESULT 7

```
US-09-925-301-505
; Sequence 505, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
```

```
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 505
; LENGTH: 1989
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1917)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-505
```

```
Alignment Scores: 2.82e-171 Length: 1989
Pred. No.: 1460.50 Matches: 277
Score: 1460.50 Conservative: 36
Percent Similarity: 88.42% Mismatches: 35
Best Local Similarity: 78.25% Indels: 6
Query Match: 80.20% Gaps: 1
DB:
```

US-10-698-228-1 (1-351) x US-09-925-301-505 (1-1989)

```
Qy 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
Db 185 CTTGAGCGGACCCCGCTCTGGCCAGCAAGCCGAGGAGGATCTTCCAGAGCCAC 244
Qy 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 245 GGAGCCGAAACTAAAGCAGCTGCCCCCGCGTGGAGGATGACCGCTGCTGAGAGAAA 304
Qy 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGl 58
Db 305 CCCCCTCGCTTTGTCTATCTTCCCATCGAGTACCATGATATCTGCGAGATGTATAAGAA 364
Qy 58 nAlaGlnAlaSerPheTrpAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
Db 365 GGCAGAGGCTTCTTTTGGACCCGCGAGGAGTGACCTCTCCAGGACATTCAGCAGTG 424
Qy 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
Db 425 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCATGTTCTGGCTTCTTTGAGC 484
Qy 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 485 AAGCGATGGCATAGTAATAAATGAAATTTGGTGGAGCGATTAGCCAAAGAAAGTTTCAGATTAC 544
Qy 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 545 AGAAGCCCGCTGTTCATGGCTTCCAAATTTGCCATGGAAACATACATTTCTGAAATGTA 604
Qy 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db 605 TAGTCTCTTATTCACACTTACATAAAGATCCCAAGAAAGGAAATTTCTTCAATGC 664
Qy 158 alleGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
Db 665 CATTGAAACGATGCTTGTGTCAAGAAAGAGCGAGACTGGGCTTGCCTGGATTGGGA 724
Qy 178 pArgLysSerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePh 198
Db 725 CAAAGAGGCTACCTTGTGAACGCTGTGTAGCCCTTTGCTCAGTGGAGGACATTTCTT 784
Qy 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db 785 TTCCGGTCTTTTGGCTCGATATTTCTGGCTCAAGAAACGAGGACTGTATGCTTGGCTCAC 844
Qy 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db 845 ATTTTCTAATGAATTAATAGCAGATGAGGGTTTACACTGTGATTTGCTTGGCTGAT 904
Qy 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db 905 GTTCAAAACCTTGTATACAAACCATCGGAGGAGAGATTAAGAGAAATAATTTATCAATGC 964
```


; PRIOR APPLICATION NUMBER: US/60/235,134
 ; PRIOR FILING DATE: 2000-09-25
 ; PRIOR APPLICATION NUMBER: US/60/235,637
 ; PRIOR FILING DATE: 2000-09-26
 ; PRIOR APPLICATION NUMBER: US/60/235,638
 ; PRIOR FILING DATE: 2000-09-26
 ; PRIOR APPLICATION NUMBER: US/60/235,711
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,720
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,840
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,863
 ; PRIOR FILING DATE: 2000-09-27
 ; NUMBER OF SEQ ID NOS: 2276
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 724
 ; LENGTH: 2500
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-954-456-724

Alignment Scores:
 Pred. No.: 4,04e-171 Length: 2500
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 9 Gaps: 1

US-10-698-228-1 (1-351) x US-09-954-456-724 (1-2500)

QY 4 ProGluArgProGluAlaAlaGluLeuAspGluArgSerSerSerAspThrAs 23
 DB 302 CCTGACGGGACCCGGCTCTGCCACAGCGGAGGAGATCTTCAGAGGCCAC 361
 QY 23 nGluSerGluLeuLysSer-----AsnGluGluProLeuLeuArgLysSe 38
 DB 362 GGAGCGGAAACTAAAGCAGCTGCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAA 421
 QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysG 58
 DB 422 CCCCCCGCGCTTGTCTATCTTCCCATCGAGTACCATGATATCTGCAGATGTATAAGAA 481
 QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
 DB 482 GGCAGAGGCTTCTTTTGGACCCCGGAGGAGGTGACCTCTCAGAGCAATTCAGCACTG 541
 QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98
 DB 542 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCCATGTTCTGGCTTTCTTGCAGC 601
 QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
 DB 602 AAGCGATGTCATAGTAATAAAGTGGTGGAGCGCATTTAGCAAGAAAGTTCCAGATTAC 661
 QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
 DB 662 AGAAGCCCGCTGTTCTATGGCTTCAAAATGGCATGCAAGAAAGGAAATTTCTTCAATGC 721
 QY 138 rSerLeuLeuLeuAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
 DB 722 TAGTCTTCTATTGACACTTACATAAAGATCCCAAGAAAGGAAATTTCTTCTCAATGC 781
 QY 158 alLeGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
 DB 782 CATTTAAACGATGCTTGTCTCAAGAAGAGGAGGAGCTGGCGCTTGCCTGGATGGGA 841
 QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
 DB 842 CAAAGAGGCTACCTATGGTGAACGTTGTAGCTTTTGTGCTGAGTGGAGGCAATTTCTTT 901
 QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218

DB 902 TTCGGTCTTTTGGCGGATATTCTGGCTCAAGAAAGAGGACTGATGCTTGGCTCAC 961
 QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 DB 962 ATTTCTTAATGAACCTATTAGCAGAGATGAGGGTTTACACTGTGATTGCTTGGCTGAT 1021
 QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluLeuIleValAspAl 258
 DB 1022 GTTCAAAACACTGCTACACAAACCATCGAGGAGAGATAGAGAAATAATATCAATGC 1081
 QY 258 aValLysIleGluGlnGluPheIleThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 DB 1082 TGTTCGATAGAACAGAGTTCTCTACTGAGGCTTGTGCTGAGGCTCATTTGGATGAA 1141
 QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG 298
 DB 1142 TTGCACCTTAATGAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACTGG 1201
 QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG 318
 DB 1202 TTTTAGCAGGTTTTCAGAGTAGAGAACCCATTGTGACTTTATGGAGATAATTTCACTGGA 1261
 QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
 DB 1262 AGGAAAGACTAATCTTTTGGAGAGAGAGTAGGCGAGTATCAGAGGATGGAGGTGATGTC 1321
 QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1322 AGTCCACAGAGAAATTTTACCTTGGATGCTGACTTC 1361

RESULT 10

; US-09-954-456-1169
 ; Sequence 1169, Application US/09954456
 ; Patent No. US20020115057A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Young, Paul
 ; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
 ; TITLE OF INVENTION: Sets
 ; FILE REFERENCE: 689290-76
 ; CURRENT APPLICATION NUMBER: US/09/954,456
 ; CURRENT FILING DATE: 2001-09-18
 ; PRIOR APPLICATION NUMBER: US/60/233,617
 ; PRIOR FILING DATE: 2000-09-18
 ; PRIOR APPLICATION NUMBER: US/60/234,052
 ; PRIOR FILING DATE: 2000-09-20
 ; PRIOR APPLICATION NUMBER: US/60/234,923
 ; PRIOR FILING DATE: 2000-09-25
 ; PRIOR APPLICATION NUMBER: US/60/235,134
 ; PRIOR FILING DATE: 2000-09-25
 ; PRIOR APPLICATION NUMBER: US/60/235,637
 ; PRIOR FILING DATE: 2000-09-26
 ; PRIOR APPLICATION NUMBER: US/60/235,638
 ; PRIOR FILING DATE: 2000-09-26
 ; PRIOR APPLICATION NUMBER: US/60/235,711
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,720
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,840
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,863
 ; PRIOR FILING DATE: 2000-09-27
 ; NUMBER OF SEQ ID NOS: 2276
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 1169
 ; LENGTH: 2500
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-954-456-1169

Alignment Scores: 4.04e-171 Length: 2500
 Pred. No.: 1460.50 Matches: 277
 Score:

Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 9 Gaps: 1

US-10-698-228-1 (1-351) x US-09-954-456-1169 (1-2500)

Qy 4 ProGluArgProGluAlaLaglyLeuAspGluArgSerSerSerAspThrAs 23
 Db 302 CCTGAGCGGACCCGGCTCTGCGCAGCAAGACCGGAGGATCTTCCAGGAGCCAC 361
 Qy 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
 Db 362 GGAGCCGAAACTAAAGCAGCTCCCGCGCGTGGAGGATGAGCGCTGCTGAGAGAAA 421
 Qy 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysG1 58
 Db 422 CCCCCCGCGCTTGTCTATCTTCCCATCGAGTACCATGATATCTGGCAGATGATATAAGAA 481
 Qy 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
 Db 482 GGCAGAGGCTTCTTTTGGACCGCCGAGGAGGTTGACCTCTCCAAGCACATTCAGCACTG 541
 Qy 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98
 Db 542 GGAATCCCTGAACCCGAGGAGAGATATTTATATCCATGTTCTTGGCTTCTTTCGAGC 601
 Qy 98 aSerAspGlyIleValAsnGluLeuValGluArgPheSerGlnGluValGlnValPr 118
 Db 602 AAGCGATGGCATAGTAATAAGAACTTGGTGGAGCGATTTAGCCAAAGATTGAGATTAC 661
 Qy 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
 Db 662 AGAAGCCCGCTGTTCTATGGCTTCCAAATGGCATGGAAACATACATTTCTGAATGTA 721
 Qy 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAla 158
 Db 722 TAGTCTCTTATTGACACTTACATAAAGATCCCAAGAAAGGAATTTCTTCTTCAATGC 781
 Qy 158 alIeGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
 Db 782 CATTTGAACGATGCCCTTGTGTCAAGAAAGAGGAGGAGCTGGGCTTGGCTGGATGGGA 841
 Qy 178 pArgLysSerThrPheGluGluArgValValAlaPheAlaAlaValGluGluValPhePh 198
 Db 842 CAAGAGGCTACCTATGGTGAAGCTGTTGATGGCTTTCGTCAGTGAAGCATTTCTT 901
 Qy 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 Db 902 TTCGGTCTTCTTGGCTCGATATTTCTGGCTCAAGAAACGAGGACTGATGCTGGCCTCAC 961
 Qy 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 Db 962 ATTTCTTAATGAATTAATAGCAGATGAGGGTTTACACTGTGATTTTGTCTGCTGAT 1021
 Qy 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAla 258
 Db 1022 GTTCAAAACCTGGTACACAAACCTCGGAGGAGAGATGAAGAAATAATATCAATGC 1081
 Qy 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 Db 1082 TGTTCCGATAGACAGAGTTCTCTCATGAGGCTTGGCTGTGAGCTCATTTGGATGAA 1141
 Qy 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG1 298
 Db 1142 TTGCACCTCTAAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAAC 1201
 Qy 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
 Db 1202 TTTTAGCAGGTTTTCAGAGTAGAACCACTTATGATTTTATGAGAGATATTTTCACTGGA 1261
 Qy 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338

Db 1262 AGGAAGACTAATCTTTTGGAGAGAGTAGGCGAGTATCAGAGATGGAGTGATGTC 1321
 Qy 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1322 AAGTCCACACAGAGAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 11

US-09-954-456-1827
 ; Sequence 1827, Application US/09954456
 ; Patent No. US20020115057A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Young, Paul
 ; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can
 ; FILE REFERENCE: 689290-76
 ; CURRENT APPLICATION NUMBER: US/09/954,456
 ; CURRENT FILING DATE: 2001-09-18
 ; PRIOR APPLICATION NUMBER: US/60/233,617
 ; PRIOR FILING DATE: 2000-09-18
 ; PRIOR APPLICATION NUMBER: US/60/234,052
 ; PRIOR FILING DATE: 2000-09-20
 ; PRIOR APPLICATION NUMBER: US/60/234,923
 ; PRIOR FILING DATE: 2000-09-25
 ; PRIOR APPLICATION NUMBER: US/60/235,134
 ; PRIOR FILING DATE: 2000-09-25
 ; PRIOR APPLICATION NUMBER: US/60/235,637
 ; PRIOR FILING DATE: 2000-09-26
 ; PRIOR APPLICATION NUMBER: US/60/235,638
 ; PRIOR FILING DATE: 2000-09-26
 ; PRIOR APPLICATION NUMBER: US/60/235,711
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,720
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,840
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,863
 ; PRIOR FILING DATE: 2000-09-27
 ; NUMBER OF SEQ ID NOS: 2276
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 1827
 ; LENGTH: 2500
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-954-456-1827

Alignment Scores:
 Pred. No.: 4,04e-171 Length: 2500
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 9 Gaps: 1

US-10-698-228-1 (1-351) x US-09-954-456-1827 (1-2500)

Qy 4 ProGluArgProGluAlaAlaGlyLeuAspGluArgSerSerSerAspThrAs 23
 Db 302 CCTGAGCGGACCCGGCTCTGCGCAGCAAGACCGGAGGATGATCTTCCAGGAGCCAC 361
 Qy 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
 Db 362 GGAGCCGAAACTAAAGCAGCTCCCGCGCGTGGAGGATGAGCGCTGCTGAGAGAAA 421
 Qy 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysG1 58
 Db 422 CCCCCCGCGCTTGTCTATCTTCCCATCGAGTACCATGATATCTGGCAGATGATATAAGAA 481
 Qy 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
 Db 482 GGCAGAGGCTTCTTTTGGACCGCGAGGAGGTTGACCTCTCCAAGCACATTCAGCACTG 541
 Qy 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98

```

Db 542 GGATCCCTGAACCCGAGAGAGATATTTATATCCATGTTCTGGCTTTCTTTCAGC 601
Qy 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 602 AAGCGATGCGCATAGTAATGAAACCTGGTGAGCGGATTTAGCCAGAAAGTTTCAGATTAC 661
Qy 118 oGluAlaArgCysPheThrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 662 AGAAGCCCGCTGTTCTATGGCTTCCAAATTTGCCATGGAAACATACATTTCTGAATGTA 721
Qy 138 rSerLeuLeuIleAspThrTyrlleArgAspProLysLysLysLysLysLysLysLysLys 158
Db 722 TAGTCTCTTATTGACACTTACATTAATAAGATCCCAAGAAGGGAATTTCTCTCAATGC 781
Qy 158 alleGluThrMetProTyrrValLysLysLysLysLysLysLysLysLysLysLysLysLys 178
Db 782 CATTTGAAACGATGCTTGTCTCAAGAAGAACGAGAGCTGGGCTTGGCTGGATGGGA 841
Qy 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
Db 842 CAAAGAGGCTACCTATGCTGAACGCTGTTGTAGGCTTTGCTGCAAGTGAAGGCATTTCTT 901
Qy 198 eSerGlySerPheAlaAlaIlePheThrLeuLysLysLysLysLysLysLysLysLysLys 218
Db 902 TTCCGGTCTTTTGGCTGCATATTTCTGCTCAAGAACGAGGAGTGTATGCTGCGCTCAC 961
Qy 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db 962 ATTTCTAATGACTTATTAGCAGAGATGAGGGTTTACACTGTGATTTGCTTGCCTGAT 1021
Qy 238 tPheGlnTyrrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db 1022 GTTCAAAACACCTGGTACACAAACCATCGGAGGAGAGATGAAGAGAATAATATTCAATGC 1081
Qy 258 aValLysIleGluGlnGluPheThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db 1082 TGTTCGGATAGAACAGAGTTCTCTACTGAGGCTTGGCTGTTGAAGCTCATTTGGGATGA 1141
Qy 278 nCysIleLeuMetLysGlnTyrrIleGluPheValAlaAspArgLeuLeuValGluLeuGl 298
Db 1142 TTGCACCTTAATGAGCAATACATTGATTTTGTGGCAGACAGACTTATGCTGGAACCTGGG 1201
Qy 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
Db 1202 TTTTAGCAAGTTTTCAGAGTAGAGAACCCATTGACTTTATGGAGAATATTTCACTGGA 1261
Qy 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrrGlnArgPheAlaValMetAl 338
Db 1262 AGGAAAGACTAACTTTTGGAGAGAGAGTAGGCGAGATATCAGAGGATGGGAGTGATGTC 1321
Qy 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1322 AGTCCAAACAGAAATCTTTTACCTTGGATGCTGACTTC 1361

```

RESULT 12

```

US-10-641-643-1370
; Sequence 1370, Application US/10641643
; Publication No. US20040077003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; Susan G. Stuart
; Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
; STREETS: 3174 PORTER DRIVE
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/641,643
; FILING DATE: 14-Aug-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1370:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2500 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g36154
; SEQUENCE DESCRIPTION: SEQ ID NO: 1370 :
US-10-641-643-1370

```

```

Alignment Scores:
Pred. No.: 4,04e-171 Length: 2500
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 18 Gaps: 1

```

```

US-10-698-228-1 (1-351) x US-10-641-643-1370 (1-2500)
Qy 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
Db 302 CCTGAGCGGAGACCGCTCTGCGCCAGCAGACCGCGAGGAGATCTTCCAGGAGCCAC 361
Qy 23 nGluSerGluLeuLysSer-----AsnGluGluProLeuLeuLeuArgLysSe 38
Db 362 GGAGCCGAAACTAAAGCAGCTGCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAAA 421
Qy 38 rSerArgArgPheValIlePhePheProfileGlnTyrrProAspIleThrLysMetTyrlsGl 58
Db 422 CCCCAGCGCTTTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481
Qy 58 nAlaGlnAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
Db 482 GGCAGAGGCTTCTTTTGGACCGCGAGGAGGTGACCTCTCCAAGGACATTCAGCAGTG 541
Qy 78 pAsnLysLeuLysAlaAspGluLysTyrrPheIleSerHisIleLeuAlaPhePheAlaAl 98
Db 542 GGAATCCCTGAACCCGAGGAGAGATATTTATATCCCATGTTCTGGCTTTCTTTCAGC 601
Qy 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 602 AAGCGATGCGCATAGTAATGAAACTTGGTGGAGCGATTAGCCAAAGAAAGTTTCAGATTAC 661
Qy 118 oGluAlaArgCysPheThrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 662 AGAAGCCCGCTGTTCTATGGCTTCCAAATTTGCCATGGAAACATACATTTCTGAATGTA 721
Qy 138 rSerLeuLeuIleAspThrTyrlleArgAspProLysLysLysLysLysLysLysLysLys 158
Db 722 TAGTCTCTTATTGACACTTACATTAATAAGATCCCAAGAAGGGAATTTCTCTCAATGC 781
Qy 158 alleGluThrMetProTyrrValLysLysLysLysLysLysLysLysLysLysLysLysLys 178

```

```

Db      782 CATTGAACGATGCTTGTCTCAAGAAAGAGCGAGACTGGCCCTTGGCTGGATGGGA 841
Qy      178 pArgLysSerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePh 198
Db      842 CAAAGAGCTACCTATGTGTGAACGTGTGTAGCCTTGTGCTGCACTGGAAGCAATTCTT 901
Qy      198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db      902 TTCGGTCTCTTTTTCGGTCGATATTCTGGCTCAAGAAACGAGGACTGATGCTGCGCTCAC 961
Qy      218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db      962 ATTTTCTAAATGAACCTATTATGACAGAGATGAGGGTTTACACTGTGATTTGCTGCCCTGAT 1021
Qy      238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAl 258
Db      1022 GTTCAACACCTGGTACACAAACCATCGAGGAGAGATGAGAGAAATATTATCAATGC 1081
Qy      258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db      1082 TGTTCGGATAGAACAGAGTTCTCTCACTGAGGCCCTTGCCTGTGAAGCTCATTTGGATGAA 1141
Qy      278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGl 298
Db      1142 TTGCACCTCTAATGAACCAATACATTTGAGTTTGTGGCAGACAGACTTATGCTGGAACCTGGG 1201
Qy      298 yPheSerLysValPheGlnAlaGluAenProPheAspPheMetGluAenIleSerLeuGl 318
Db      1202 TTTTAGCAAGGTTTTCAGATAGAACCCATTTGACTTTATGGAGAAATATTTCACTGGA 1261
Qy      318 uGlyLysThrAsnValPheGlyLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
Db      1262 AGGAAGACTAACTTTCTTGAGAGAGAGATAGCGGACTATCAGAGGATGGAGTGTGTC 1321
Qy      338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db      1322 AAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361

```

RESULT 13

```

US-10-733-878-458
; Sequence 458, Application US/10733878
; Publication No. US20040224408A1
; GENERAL INFORMATION:
; APPLICANT: Jean-Philippe Girard
; APPLICANT: Myriam Rousseigne
; APPLICANT: Francois Amalric
; TITLE OF INVENTION: THAP PROTEINS AS NUCLEAR RECEPTORS FOR
; TITLE OF INVENTION: CHEMOKINES AND ROLES IN TRANSCRIPTIONAL REGULATION, CELL
; TITLE OF INVENTION: PROLIFERATION AND CELL DIFFERENTIATION
; FILE REFERENCE: BIOBANK.012A
; CURRENT APPLICATION NUMBER: US/10733,878
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 60/432699
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/485027
; PRIOR FILING DATE: 2003-07-03
; NUMBER OF SEQ ID NOS: 535
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 458
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-733-878-458

Alignment Scores:
Pred. No.: 4,04e-171 Length: 2500
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
Gaps: 20

```

```

US-10-698-228-1 (1-351) x US-10-733-878-458 (1-2500)
Qy      4 ProGluArgProGluAlaAlaGlyLeuAspGlu-ArgSerSerSerAspThrAs 23
Db      302 CCTGAGCGGACCCCGCTCTGGCCAGCAAGACCCGAGGAGGATCTTCCAGGAGCCAC 361
Qy      23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db      362 GGAGCCGAACTAAAGCAGCATGCGCCCGCGGTGGAGGATGAGCGCTCTCTGAGAGAAA 421
Qy      38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGl 58
Db      422 CCCCCCGCGCTTGTGCATCTTCCCATCGAGTACCATGATATCTGCAGCATGTATAGAA 481
Qy      58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
Db      482 GGCAGAGCTTCTTCTTGGACCGCGAGGAGGTTCACCTCTCCAGGACATTCAGCACTG 541
Qy      78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
Db      542 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCCATGTTCTGGCTTCTTTTCGAGC 601
Qy      98 aSerAspGlyIleValAsnGluAenLeuValGluArgPheSerGlnGluValGlnValPr 118
Db      602 AAGCGATGSCATAGTAAATGAAACTTGTGTGGAGCGATTTAGCCAAAGAAATTCAGATTAC 661
Qy      118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db      662 AGAAGCCCGCTTGTCTATGGCTTCCAAATTCGCAATGGAAACATACATCTTGAAATGT 721
Qy      138 rSerLeuLeuIleAspThrTyrIleAArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db      722 TAGTCTTCTTATGACACTTACATAAAGATCCCAAGAAAGGAAATTTCTCTTCAATGC 781
Qy      158 aIleGluThrMetProTyrValLysLysIleAlaAspTrpAlaLeuArgTrpIleAlaAs 178
Db      782 CATTGAAACGATGCTTGTGTCAAGAGAAAGCAGACTGGCGCTTGGCTGGATTGGGGA 841
Qy      178 pArgLysSerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePh 198
Db      842 CAAAGAGCTACCTATGTGTGAACGTGTGTAGCCTTGTGCTGAGTGAAGGCAATTTCTT 901
Qy      198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db      902 TTCGGTCTCTTTTTCGGTCGATATTCTGGCTCAAGAAACGAGGACTGATGCTGGCTCAC 961
Qy      218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db      962 ATTTTCTAATGAACCTATTAGCAGAGATGAGGGTTTACACTGTGATTTTGTTCCTGAT 1021
Qy      238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAl 258
Db      1022 GTTCAACACCTGGTACACAAACCATCGAGGAGAGATGAGAGAAATATTATCAATGC 1081
Qy      258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db      1082 TGTTCGGATAGAACAGGAGTTCTCTCTGAGGCCCTTGTGCTGAGCTCATTTGGGATGAA 1141
Qy      278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGl 298
Db      1142 TTGCACCTCTAATGAACCAATACATTTGAGTTTGTGGCAGACAGACTTATGCTGGGATGAA 1141
Qy      298 yPheSerLysValPheGlnAlaGluAenProPheAspPheMetGluAenIleSerLeuGl 318
Db      1202 TTTTAGCAAGGTTTTCAGATAGAACCCATTTGACTTTATGGAGAAATATTTCACTGGA 1261
Qy      318 uGlyLysThrAsnValPheGlyLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
Db      1262 AGGAAGACTAACTTTCTTGAGAGAGAGATAGCGGACTATCAGAGGATGGAGTGTGTC 1321
Qy      338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351

```

Db 1322 AAGTCCACAGAGAAATCTTTACCTTGGATGCTGACTTC 1361

RESULT 14

US-10-843-641A-3751
 ; Sequence 3751, Application US/10843641A
 ; Publication No. US20050064454A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Avalon Pharmaceuticals, Inc.
 ; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
 ; FILE OF INVENTION: Signature Gene Sets
 ; FILE REFERENCE: 689290-189
 ; CURRENT APPLICATION NUMBER: US/10/843,641A
 ; CURRENT FILING DATE: 2004-05-12
 ; PRIOR APPLICATION NUMBER: US/09/873,367
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: US/09/954,531
 ; PRIOR FILING DATE: 2001-09-18
 ; PRIOR APPLICATION NUMBER: US/09/954,456
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US/09/962,436
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US/09/962,832
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US/09/964,824
 ; PRIOR FILING DATE: 2001-09-27
 ; PRIOR APPLICATION NUMBER: US/09/967,768
 ; PRIOR FILING DATE: 2001-09-28
 ; PRIOR APPLICATION NUMBER: US/09/968,007
 ; PRIOR FILING DATE: 2001-10-02
 ; PRIOR APPLICATION NUMBER: US/09/969,347
 ; PRIOR FILING DATE: 2001-10-02
 ; PRIOR APPLICATION NUMBER: US/09/969,708
 ; PRIOR FILING DATE: 2001-10-03
 ; Remaining prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 8447
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 3751
 ; LENGTH: 2500
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-843-641A-3751

Alignment Scores:
 Pred. No.: 4,04e-171 Length: 2500
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 21 Gaps: 1

US-10-698-228-1 (1-351) x US-10-843-641A-3751 (1-2500)

Qy 4 ProGluArgProGluAlaAlaGlyLeuAspGluAspGlu-ArgSerSerSerAspThrAs 23
 Db 302 CTGAGCGGAGCCCGCGTCTCTGCCACAGACCGCGGAGGAGGATCTTCCAGGAGCCAC 361
 Qy 23 nGluSerGluLeuLysSer-----AsnGluGluProLeuLeuArgLysSe 38
 Db 362 GGAGCGGAAACTAAGACGAGCTGCCCGCGGAGGAGGAGCGCTGCTGAGAGAA 421
 Qy 38 rSerArgArgPheValIlePheProIlePheProIlePheProIlePheProIlePhe 58
 Db 422 CCGCGCGCGCTTGTTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGATAAGAA 481
 Qy 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
 Db 482 GGCAGAGGCTTCTTTTGGACCGCGGAGGAGGTTGACCTCTCCAGGACATTCAGCACTG 541
 Qy 78 pAsnLysLeuLysAlaAspGluLysTrpPheIleSerHisIleLeuAlaPhePheAla 98
 Db 542 GGAATCCCTGAAACCCGAGGAGGATATTATATCCATGTTCTGCTTCTTTCGAGC 601
 Qy 98 aSerAspGlyIleValAsnGluLeuValGluArgPheSerGlnGluValGlnValPr 118

Db 602 AAGCGATGGCATAGTAATGAAACCTTGGTGGCGGATTTAGCCAGAGTTCAATTAC 661
 Qy 118 oGluAlaArgCysPheTrpGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
 Db 662 AGAAGCCCGCTGTTTCTATGGCTTCCAATTTGCCATGGAAACATACATCTCTGAATGTA 721
 Qy 138 rSerLeuLeuIleAspThrTyIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
 Db 722 TAGTCTTCTTATTGACACTTACATAAAAGATCCCAAGAAAGGAATTTCTCTCAATGC 781
 Qy 158 aIleGluThrMetProTyTrpValLysLysLeuAlaAspTrpAlaLeuArgTrpIleAlaAs 178
 Db 782 CATTGAAACCATGCTTGTGTCAAGAAAGAGCGAGACTGGGCGCTTGCCTGGAATGGGA 841
 Qy 178 pArgLysSerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePh 198
 Db 842 CAAAGAGGCTACCTATGCTGACCGTGTGTAGCCCTTGTCTGAGTGGAGGCAATTTCTT 901
 Qy 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 Db 902 TTCGGTCTTCTTTCGTCGATATTCTGGCTCAAGAAACGAGGACTGATGCTGCGCTCAC 961
 Qy 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 Db 962 ATTTTCTTAATGAACCTTATTAGCAGATGAGGGTTTACACTGTGATTTTGTCTGCTGAT 1021
 Qy 238 tPheGlnTyLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
 Db 1022 GTTCAACACCTGGTACACAAACCATCGGAGGAGAGTAGAGAAATATATATCAATGC 1081
 Qy 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 Db 1082 TGTTCGATAGAACAGGAGTTCTCTCACTGAGGCTTGCCTGGAAGCTCATTTGGGATGAA 1141
 Qy 278 nCysIleLeuMetLysGlnTyIleGluPheValAlaAspArgLeuLeuValGluLeuGl 298
 Db 1142 TTGCACCTTAATGAACAAATACATTTGATTTGTGGCAGACAGACTTATGCTGGAACCTGG 1201
 Qy 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
 Db 1202 TTTTAGCAGGTTTTCAGAGTAGAGAACCCCATTTGACTTTATGGAGAAATATTTCACTGA 1261
 Qy 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyTrpGlnArgPheAlaValMetAl 338
 Db 1262 AGGAAAGACTAATCTTTTCAGAGAGAGTAGGCGAGTATCAGAGGATGGGAGTGCATGTC 1321
 Qy 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1322 AAGTCCACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 15

US-10-843-641A-4196
 ; Sequence 4196, Application US/10843641A
 ; Publication No. US20050064454A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Avalon Pharmaceuticals, Inc.
 ; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
 ; FILE OF INVENTION: Signature Gene Sets
 ; FILE REFERENCE: 689290-189
 ; CURRENT APPLICATION NUMBER: US/10/843,641A
 ; CURRENT FILING DATE: 2004-05-12
 ; PRIOR APPLICATION NUMBER: US/09/873,367
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: US/09/954,531
 ; PRIOR FILING DATE: 2001-09-18
 ; PRIOR APPLICATION NUMBER: US/09/954,456
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US/09/962,436
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US/09/962,832
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US/09/964,824

```
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4196
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-4196

Alignment Scores:
Pred. No.: 4,04e-171 Length: 2500
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 21 Gaps: 1

US-10-698-228-1 (1-351) x US-10-843-641A-4196 (1-2500)

QY 4 ProGluArgProGluAlaAlaGlyLeuAspGluArgSerSerSerAspThrAs 23
DB |||||
QY 302 CCTGAGCGGAGCCCGGCTCTGGCCAGCAAGACCGCGAGGAGGATCTCCAGGAGCCAC 361
DB |||||
QY 23 nGluSerGluLeuLysSer-----AsnGluGluProLeuLeuArgLysSe 38
DB |||||
QY 362 GGAGCGGAACCTAAAGCAGCTGCCCGCGGTGGAGGATGAGCGGTCTGAGAGAAA 421
DB |||||
QY 38 rSerArgArgPheValIlePhePheProIleGlnTyrProAspIleTyrLysMetTyrLysG 58
DB :|||||
QY 422 CCCCCCGCGCTTTGTCTATCTCCCATCGAGTACCATGATATCTGGCAGATGATATAAGAA 481
DB |||||
QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
DB :|||||
QY 482 GGCAGAGGCTTCCTTTTGGACCGCGGAGGAGGTGACCTCTCCAAGACATTCAGCACTG 541
DB |||||
QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAla 98
DB |||||
QY 542 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCATGTTCTGGCTTCTTTGCGAGC 601
DB |||||
QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
DB |||||
QY 602 AAGGATGGCATAGTAAATGAAACCTTGGTGGAGCGATTAGCCAGAGATTCAGATTAC 661
DB |||||
QY 118 oGluAlaArgCyPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
DB |||||
QY 662 AGAAGCCCGCTGTTTCTATGGCTTCAAATGGCATGCAAAACATACATCTGAAATGTA 721
DB |||||
QY 138 rSerLeuLeuLeuAspThrTyrIleArgAspProLysArgGluPheLeuPheAenAl 158
DB |||||
QY 722 TAGCTCTCTATTGACACTTACATAAAGATCCAAAGAGGGAATTTCTCTTCAATGC 781
DB |||||
QY 158 aIleGluThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
DB |||||
QY 782 CATTGAAACGATGCTTGTCTCAGAGAGAGGAGGAGCTGGCGCTTGGCTGGATTGGGGA 841
DB |||||
QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
DB |||||
QY 842 CAAAGAGGCTACCTATGGTCAAAGTGTGTAGCTTTGCTGCTGAGTGAAGGCATTTTCTT 901
DB |||||
QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
DB |||||
QY 902 TTCGGTCTCTTTTGGCTCGATATTTCTGGCTCAAGAAACGAGGAGCTGATGCTGGCTCAC 961
DB |||||
QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
DB |||||
```

Search completed: June 16, 2005, 18:52:15
Job time : 722 secs

```
DB 962 ATTTTCTAATGAACCTTATTAGCAGAGATGAGGTTTACCTGTGATTTTGTCTGCTGAT 1021
QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
DB |||||
QY 1022 GTTCAAAACACCTGGTACACAAACCATCGGAGGAGAGTAAGAGAAATAATTATCAATGC 1081
DB |||||
QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
DB |||||
QY 1082 TGTTCGGATAGAACAGGAGTTCTCTCAGTGGCCCTTGGCTGTGAAGCTCATTTGGGATGAA 1141
DB |||||
QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuG 298
DB |||||
QY 1142 TTGCACCTCTAATGAAGCAATACATTGAGTTTGTGGAGACAGACTTATCTGGAAGCTGGG 1201
DB |||||
QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG 318
DB |||||
QY 1202 TTTTAGCAGGTTTTCAGAGTAGAGAACCCATTTCACCTTATGGAGAAATATTTCACTGGA 1261
DB |||||
QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
DB |||||
QY 1262 AGGAAAGAGACTAACTTCTTTGAGAGAGAGTAGGCGAGTATCAGAGGATCGGAGTGATGTC 1321
DB |||||
QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB :|||||
QY 1322 AAGTCCAACAGAGAAATCTTTTACCTGGATGCTGACTTC 1361
DB |||||
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 16, 2005, 11:03:33 ; Search time 3856 Seconds
(without alignments)

3464.875 Million cell updates/sec

Title: US-10-698-228-1

Perfect score: 1821

Sequence: 1 MGDPRPEAAGLDQDERSSS.....QRPVMAETTDNVFTLDADF 351

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Delop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-frame_p2n.model -DEV=xlh

-Q/cgna2_1/USPTO.spool/US10698228/runat_14062005_161416_28309/app_query.fasta_1.519

-DB=EST -QFMT=fastap -SUFFIX=rest -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0

-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45

-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTFMT=pco -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US10698228_@CGN_1_1_4352_@runat_14062005_161416_28309 -NCPU=6 -ICPU=3

-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:**

1: gb_est1:**

2: gb_est2:**

3: gb_hc:**

4: gb_est3:**

5: gb_est4:**

6: gb_est5:**

7: gb_est6:**

8: gb_gse1:**

9: gb_gse2:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	100.0	3292	3	CR617553 full-leng
2	1821	100.0	4748	3	AL137348 Homo sapi
3	1791	98.4	4650	3	BC042948 Homo sapi
4	1737.5	95.4	3501	3	CR627376 Homo sapi
5	1591	87.4	1083	1	AL547501 AL547501
6	1585	87.0	1056	9	AY398973 Homo sapi
7	1468	80.6	1043	9	AY398975 Mus muscu
8	1460.5	80.2	1571	3	CR603461 full-leng
9	1460.5	80.2	1573	3	CR625489 full-leng

10	1460.5	80.2	1582	3	CR608076	full-leng
11	1460.5	80.2	1588	3	CR602054	full-leng
12	1460.5	80.2	1592	3	CR621427	full-leng
13	1460.5	80.2	1600	3	CR604378	full-leng
14	1460.5	80.2	1605	3	CR590959	full-leng
15	1460.5	80.2	1612	3	CR614990	full-leng
16	1460.5	80.2	1613	3	CR609838	full-leng
17	1460.5	80.2	1623	3	CR618451	full-leng
18	1460.5	80.2	1630	3	CR602150	full-leng
19	1460.5	80.2	1796	3	CR596700	full-leng
20	1443	79.2	2088	3	AK088907	Mus muscu
21	1384	76.0	879	5	BQ441857	AGENCOURT
22	1354	74.4	892	7	CNI63214	952695 MA
23	1328.5	73.0	997	4	BM468712	AGENCOURT
24	1316	72.3	1008	9	AY398974	Pan trogl
25	1277	70.1	932	5	BUI196941	AGENCOURT
26	1275	70.0	914	7	CF995079	AGENCOURT
27	1257.5	69.1	946	7	CN024619	AGENCOURT
28	1235	67.8	870	7	CO648380	ILLUMIGEN
29	1222	67.1	836	7	CR416843	CR416843
30	1221	67.1	893	5	BUI190680	AGENCOURT
31	1209.5	66.4	923	7	CR580780	CR580780
32	1209	66.4	1076	4	BM460735	AGENCOURT
33	1203	66.1	801	5	BP680185	BP680185
34	1197.5	65.8	933	6	CA983633	AGENCOURT
35	1195	65.6	1111	5	BM914217	AGENCOURT
36	1189	65.3	909	6	CA981614	AGENCOURT
37	1188	65.2	804	6	CD656661	AGENCOURT
38	1186.5	65.2	1028	5	BQ050629	AGENCOURT
39	1186	65.1	879	5	BUI189053	AGENCOURT
40	1186	65.1	910	6	CD794331	EST665692
41	1177	64.6	909	7	CK179258	EST768578
42	1170	64.3	928	4	BM018808	603646723
43	1168.5	64.2	1080	6	CD461172	FG09_04p0
44	1167	64.1	896	5	BUI170979	AGENCOURT
45	1164	63.9	838	5	BUI10710	AGENCOURT

ALIGNMENTS

RESULT 1	CR617553	3292 bp	mrna	linear	HTC 21-JUL-2004
LOCUS	full-length cDNA clone CSODI011YF14 of Placentia Cot 25-normalized				
DEFINITION	full-length cDNA clone CSODI011YF14 of Placentia Cot 25-normalized				
ACCESSION	CR617553				
VERSION	CR617553.1				
KEYWORDS	HTC; CNSLT cDNA.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Li, W.B., Gruber, C., Jesses, J. and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished				
REMARK	Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue				
REFERENCE	2 (bases 1 to 3292)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)				
COMMENT	- Web : www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.				
FEATURES	Location/Qualifiers				
source	1..3292				
	/organism="Homo sapiens"				

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS01011yF14"
/tissue type="placenta Cot 25-normalized"
/plasmid="pCWSPORT_6"

ORIGIN

Alignment Scores:

Pred. No.: 2,61e-210 Length: 3292
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-698-228-1 (1-351) x CR617553 (1-3292)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGluArgSerSerSer 20
DB 27 ATGGGCGACCCCGAAAGCGCGAAGCGCGCGGCTGGATCAGATCAGATCATCTTCA 86
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
DB 87 GACACCAACGAAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 146
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
DB 147 CGGTTGTTCATCTTCCATCCAGTACCTCGATATTTGGAAATGTTATTAACAGGCACAG 206
QY 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
DB 207 GCTTCCTTCGACAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAACAAG 266
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaLaSerAsp 100
DB 267 CTTAAGCAGATGAGAAAGTACTTCACTCTCACTCTAGCCCTTTTTCGAGCCAGTGAT 326
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 327 GGAATTTGAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTCGAGGTTCCAGAGGCT 386
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 387 CGCTGTTTCTATGGCTTCAAAATCTCATCGAATGTTCACTCAGAGATGTACAGTTTG 446
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 447 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 506
QY 161 ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
DB 507 ACCATGCCCTATGTTAAGAAAGAGAGATGGCCCTTCGATGGATAGCAGATAGAAA 566
QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
DB 567 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGAGAGAGGATTTCTTCTCAGGA 626
QY 201 SerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 627 TCTTTTGGCTATATCTTGGCTTAAAGAGAGAGGCTTATGCGAGGACTCACTTTTTC 686
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 687 ATGAAGTCTATCAGCAGATGAGGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 746
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleLeuValAspAlaValLys 260
DB 747 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 806
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 807 ATTGACAGGAGTGTAAACAGAGGCTTCCAGGATTTGGCCTCATTTGGAATGAAATTGCAT 866

QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 867 TTGTAACACAGTACATTGAGTTGTAGCTTGACAGATTTACTTTGGAACTTTGGATCTCA 926
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 927 AAGTTTTTCAGCAGAGAAATCTTTTGAATTTATGGAACATTTCTTTAGAGAGAAA 986
QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 987 ACAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGCGAGATTTGCGAGAAACC 1046
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1047 ACAGATAACGCTCTTCACCTTGGATGCGATTTT 1079

RESULT 2
HSM802033
LOCUS Homo sapiens mRNA; cDNA DKFZp761E1312 (from clone DKFZp761E1312).
ACCESSION AL137348
VERSION AL137348.1 GI:6807859
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4748)
AUTHORS Ansorge W., Krieger S., Regiert T., Rittmueller C., Schwager B., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M. and Wiemann S.
CONSRMT The German cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp761E1312) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp761E1312
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.

FEATURES
Location/Qualifiers
1..4748
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp761E1312"
/db_xref="taxon:9606"
/clone="DKFZp761E1312"
/tissue type="amygdala"
/clone_lib="761 (synonym: hamy2). Vector pSport1; host DH10B; sites NotI + SalI"
/dev stage="adult"
/note="ribonucleotide reductase"
1..4748
/genes="DKFZp761E1312"
47..1102
/gene="DKFZp761E1312"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB70703.2"
/db_xref="GI:52545612"
/db_xref="GOA:Q9NTD8"
/translation="MGDPPEAAGLDQDERSSDTNESEIKSNEEPILLKSSRRFVI
FPQYDINWYKQAQASFTAEVDLSKDLPHNNKLKADSKYFISHILAFPAASDGI
VNSNLFVRSQEVQVPCARFYFQILINENHSEMYSLIDITYINDPKRFLFNALIE
TMPYVKKADWALRIADRKSTFGFVVAFAAEGVFVFGSFAIFWLKKGLMPGLT
FSNELISRDEGLHCDFACLMFQVILVNPSPSERVEREIIIVDAVKIEQEFTEALPVGLIG

MNCILMKQYIEFVADRLVLLVGLSPKVPQAEHPDFMENISLEGKTNPFEXRVSRYQRP
AVMAETTDNVFTLDAF*

ORIGIN

Alignment Scores:
Pred. No.: 4,56e-210 Length: 4748
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-698-228-1 (1-351) x HSM802033 (1-4748)

Qy 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 47 ATGGCGCACCAGAAAGCGCGGAGCGCGCGGCTGGATCAGGATGAGATCATCTTCA 106
Qy 21 AspThrAnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
Db 107 GACACCAACGAAAGTGAATCAATCAATCAAGAGCCACTCTCTAAGAAAGAGTCTCGC 166
Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
Db 167 CGGTTTGTCTATCTTCCAAATCCAGTACCCCTGATATTTGGAAATGTATAACAGGCACAG 226
Qy 61 AlaSerPheTyrThrLaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
Db 227 GCTTCTCTTGGACAGAGAGGTGCTGATTCATCAAGGATCTCCCTCACTGGAACAG 286
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAspAsp 100
Db 287 CTTAAAGCAGATGAGAAGTACTTCTCATCTTCACATCTTAGCCTTTTGGCAGCGAGTAT 346
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 347 GGAATTTGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGTCAGAGGCT 406
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 407 CGCTGTTTCTATGCTTTCAAATCTCATCGAATGTTTCACTCAGAGATGTACAGTTG 466
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 467 CTGATAGACATTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 526
Qy 161 ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
Db 527 ACCATGCCCTATGTTAAGAAAAAGCAGATGGCCCTTGCATGCATGACAGATAGAAA 586
Qy 181 SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 200
Db 587 TCTACTTTTGGGAAAGAGTGGTGCCCTTTGCTGTGTAGAAGAGATTTTCTTCTCAGGA 646
Qy 201 SerPheAlaAlaIlePheTyrLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 647 TCTTTTGTCTATTAATCTGCTTAAGAAAGAGAGTCTTATGCCAGGACTCACATTTTCC 706
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 707 AATGNACTCATCAGCAGATGAAGACTTCACTGTGACTTTTGTCTGCTGATGTTCCTCAA 766
Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 767 TACTTAGTAATAAGCCCTCAGAAAGAAAGGCTCAGGAGATCATTTGTGTGCTGTCAA 826
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 827 ATTGAGCAGAGGTTTTTAACAGAGCCCTTCCAGTGTGCCCTCATTTGGAATGAATTCATT 886
Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 887 TTGATGAACAGTACATTGATTGTTGTAGCTGACAGATTACTTGTGGAACTTGGATTCTCA 946

Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 947 AAGGTTTTTTCAGGCAGAAAAATCCTTTTCATTTTATGAAAAATCTTTTAGAAGAAAA 1006
Qy 321 ThrAnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 1007 ACAAATTTTCTTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1066
Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1067 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1099

RESULT 3

BC042948 4650 bp mRNA linear HTC 12-OCT-2004
LOCUS BC042948
DEFINITION Homo sapiens cDNA clone IMAGE:4798175, containing frame-shift errors.

ACCESSION BC042948
VERSION BC042948.2 GI:34194000
KEYWORDS HTC.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 4650)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butler, A.D., Schein, J.E., Jones, S.J., Skalska, U., Smal, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 4650)

Director MGC Project.

Direct Submission

Submitted (09-JAN-2003) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.ncbi.nlm.nih.gov>

On Aug 25, 2003 this sequence version replaced gi:27695575.

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca

Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,

Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth

Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,

Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,

Teika Olson, Diana Plasson, Anca Petrescu, Anna Liisa Prabh,

Parvaneh Saedi, JR Santos, Angelique Schnerch, Ursula Skalska,

Duane Smal, Jeff Stott, Miranda Tsai, George Yang, Jacque

Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAK Plate: 70 Row: h Column: 19
This clone has the following problem: frame shifted.

FEATURES

source

Location/Qualifiers
1. .4650
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone=IMAGE:4798175"
/issue_type="Brain, hypothalamus"
/clone_lib="NIH MGC_96"
/lab_host="DH10B"
/note="Vector: pBluescript"

ORIGIN

Alignment Scores:

Pred. No.: 2,04e-206 Length: 4650
Score: 1791.00 Matches: 350
Percent Similarity: 99.72% Conservatives: 0
Best Local Similarity: 99.72% Mismatches: 1
Query Match: 98.35% Indels: 1
DB: 3 Gaps: 0

US-10-698-228-1 (1-351) x BC042948 (1-4650)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
DB 72 ATGGCGGACCCGGAAGCGCGGCTGGATCAGATGAGATCATCTTCA 131
QY 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
DB 132 GACACCAAGCAAGTAAATGAATGAAGACA-CTCTAAGAAAGATTCCTGC 190
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
DB 191 CGTTTGTTCATCTTCCATCCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG 250
QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
DB 251 GCTTCTCTTGACAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAACAAG 310
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
DB 311 CTAAAGCAGATGAGAGTACTTCACTCTCACTCTTAGCCCTTTTTCAGCCAGTGTAT 370
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 371 GGAATGTAAATGAATTTTGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 430
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 431 CGCTGTTTCTATGGCTTTCAATTCATCATCGAGAATGTTCACTCAGACAGATGTACAGTTG 490
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 491 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTAATGCAATGTAA 550
QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
DB 551 ACATGCCCTATGTTTANGAANAAGACAGATTTGGCCCTTGGCATGATGACGATAGANAA 610
QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
DB 611 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGGAGTTCCTTCTCAGGA 670
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 671 TCTTTTGTGCTATATCTGGCTAAAGAGAGAGGCTTTATGCGCAGGACTCACTTTTTC 730
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240

DB 731 AATGAACCTCATCAGCAGAGATGAAGGACTTCATGTGACTTTGCTTGCTCATGTTCCAA 790
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB 791 TACTTAGTAATAAGCTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAA 850
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 851 ATTGAGCAGGAGTTTAAACAGAAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATGCATT 910
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 911 TTGATGAACACAGTACATGAGTTTGTAGCTGCACAGATTACTTTGGAACTTTGGATTCTCA 970
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 971 AAGTTTTTCAGCAGAGAAATCCTTTTGATTTTATGGAACATTTCTTTAGAGAGAAA 1030
QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 1031 ACAATTTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1090
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1091 ACAGATAACGTTCTTCCCTTTGGATGCAGATTTT 1123
RESULT 4
CR627376 3501 bp mRNA linear HTC 22-SEP-2004
LOCUS Homo sapiens mRNA; cDNA DKFZp686M05248 (from clone DKFZp686M05248).
DEFINITION CR627376
ACCESSION CR627376.1 GI:50949847
VERSION
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3501)
AUTHORS Ortenwaelder,B., Obermaier,B., Deutschenbaue,S., Schaipp,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
CONSORTM The German cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
COMMENT Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix (Martinsried/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp686M05248) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
<http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp686M05248>
Further information about the clone and the sequencing project is available at <http://mips.gsf.de/projects/cdna/>.
FEATURES
source
1. .3501
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686M05248"
/tissue_type="salivary gland"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
/note="ribonucleotide reductase M2 B (TP53 inducible)"
1. .3501
/gene="DKFZp686M05248"
332. .1225
/genes="DKFZp686M05248"
/codon_start=1
/product="hypothetical protein"
gene
CDS

```

/protein_id="CAH10473.1"
/db_xref="GI:50949848"
/translation="MYKQASFWTAAEEVDLSKDLPHNNKLKADKEYFIHILAFPAF
SDGIVNENLYVERFSQEQVPEARCFYGFQILINVENHSEMYSLIIDTYIRDPKKREFLF
NATETWPYVKKADWALRWIADKSTFGERVAFAAVEGVFFSGSPAAIIFWLKKRGLM
PGITFNNELISDRGELHCDPACILFQVKNKPEERVREIIVDAVKLEQSEFLTEALPV
GLIGMCILMKQYIEFFVADRLLYVELGFSKVFQAEINPDPFENISLEKNTFFEKRVSE
YQRFVNAWETTDINDVFLDADF"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 4 5e-200 Length: 3501
Score: 1737.50 Matches: 339
Percent Similarity: 96.06% Conservatives: 2
Best Local Similarity: 95.49% Mismatches: 7
Query Match: 95.41% Indels: 7
DB: 3 Gaps: 1

US-10-698-228-1 (1-351) x CR627376 (1-3501)

QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu----- 16
Db 158 CCCGCTCGCTGCTGCCCCGGGGCCAGCCAGCAGCGCTGGAGCGCTTCAGCCGGG 217
QY 17 ArgSerSerAspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArg 36
Db 218 AGATCATCTTCAGACACCAACGAAAGTGAATAAAGTCAATGAAGCCACTCCTAAGA 277
QY 37 LysSerSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyr 56
Db 278 AAGAGTTCGCGCGTGTGTCATCTTCCCAATCAGTACCCTGATATTTGGAAATATAT 337
QY 57 LysGlnAlaGlnAlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuPro 76
Db 338 AAACAGCAGCGCTTCTTCTGACAGCAGAGGTCGACTTATCAAGAGTCTCCCT 397
QY 77 HisTrpAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePhe 96
Db 398 CACTGGAACAGCTTAAAGCAGATGAGAGTACTTCTCATCTCTACATCTTAGCCTTTT 457
QY 97 AlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGln 116
Db 458 GCAGCCAGTGATGAAATGTAAATGAAATTTGGTGAACGCTTTAGTCAGAGGTCAG 517
QY 117 ValProGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGlu 136
Db 518 GTTCCAGAGCGCTCGCTGTTCTATGGCTTTCAAAATCTCATCAGAGATGTTCACTCAG 577
QY 137 MetTyrSerLeuLeuLeuAspThrTyrIleArgAspProLysLysArgGluPheLeuPhe 156
Db 578 ATGTACAGTTTGTGTAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTT 637
QY 157 AsnAlaIleGluThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIle 176
Db 638 AATGCAATTGAAACCATGCTTATTTAGAAAAAAGCAGATTGGCGCTTGGCATGGATA 697
QY 177 AlaAspArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluVal 196
Db 698 GCAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCCCTTCTGCTGTAGAGGAGTT 757
QY 197 PhePheSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGly 216
Db 758 TTCTTCTCAGGATCTTTGCTGCTATATCTTGCTAAAGAGAGAGGCTTATGCCAGGA 817
QY 217 LeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCys 236
Db 818 CTCACATTTTCCCAATGAATCATCAGCAGAGATGAAGCACTTCACTGTGACTTTGCTTGC 877
QY 237 LeuMetPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleVal 256
Db 878 CTGATGTTCCAAATCTAGTAAATAAGCTTCCAGAGAAAGGGTCAGGAGATCATTTGTT 937
QY 257 AspAlaValLysIleGluGlnGluPheLeuThrGlnAlaLeuProValGlyLeuIleGly 276

```

```

Db 938 GATGCTGTCAAAATGACAGAGATTTTAAACAGAGCCTTCCAGTTGGCCTCATTTGA 997
QY 277 MetAsnCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGlu 296
Db 998 ATGAATTGCAATTTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAA 1057
QY 297 LeuGlyPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSer 316
Db 1058 CTTGGATTCTCAAAGGTTTTTTCAGCAGAAAAATCCTTTTATGATTTATGAAAAATTTCT 1117
QY 317 LeuGluGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaVal 336
Db 1118 TTAGAAGGAAAAACAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTT 1177
QY 337 MetAlaGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1178 ATGCAGAAACACACAGATACGCTTTCACCTTGGATGCAGATTTT 1222

RESULT 5
AL547501 1083 bp mRNA linear EST 25-MAR-2004
LOCUS
DEFINITION AL547501 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1011YF14 5-PRIME, mRNA sequence.
ACCESSION AL547501
VERSION AL547501.3 GI:45747948
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1083)
AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:31269332.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4436.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0D1011DC07QPI&c=4436.r.

FEATURES
Source
1. 1083
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1011YF14"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized."

Alignment Scores:
Pred. No.: 5.81e-183 Length: 1083
Score: 1591.00 Matches: 319
Percent Similarity: 97.85% Conservatives: 0
Best Local Similarity: 97.85% Mismatches: 5
Query Match: 87.37% Indels: 4
DB: 1 Gaps: 0

US-10-698-228-1 (1-351) x AL547501 (1-1083)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSe 20

```

```
Db 27 ATGGCGAGCCCGGAAGCGCGGCGGCTGGATCAGATGAGAGATCATCTTC 86
Qy 20 rAspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerAr 40
Db 87 AGACACCAACGAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCG 146
Qy 40 gArgPheValIlePheProIleGlnTyrProAspIleTyrPheLysMetTyrLysGlnAlaGl 60
Db 147 CCGGTTTGTCATCTTTCCCAATCAGTACCTCGATATTTGGAAAATGTATAAACAGGCACA 206
Qy 60 nAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrPheSerLys 80
Db 207 GGCCTTCCTTCGACAGCAGAGAGTGCATTTATCAAAAGGATCTCCCTCAGTCGGAACAA 266
Qy 80 sLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAs 100
Db 267 GCTTAAGCAGATGAGAAGTACTTTCATCTCTCATCTTAGCCTTTTTCAGCCAGTGA 326
Qy 100 pGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAl 120
Db 327 TGGAAATTGTAATGAAAATTTGGTGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGC 386
Qy 120 aArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLe 140
Db 387 TCGCTGTTCTATGGCTTTCAAAATCTCATCGAATGTTCACTCAGAGATGTACAGTTT 446
Qy 140 uLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGl 160
Db 447 GCTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATCAATTGA 506
Qy 160 uThrMetProTyrValLysLysLysAlaAspTyrPalaLeuArgTyrPileAlaAspArgLy 180
Db 507 AACCATGCCCTATGTAAGAAAAAGCAGATTTGGCGCTTGGCATGATAGCAGATAGAAA 566
Qy 180 sSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGluValPhePheSerGl 200
Db 567 ATCTACTTTTGGGAAGAGTGGTGGCTTTTGTCTGTAGAGGAGTTTCTTCTCAGG 626
Qy 200 ySerPheAlaAlaIlePheTyrPheLysLysArgGlyLeuMetProGlyLeuThrPheSe 220
Db 627 ATCTTTTGTCTATATTTCTGGCTAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 686
Qy 220 rAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGl 240
Db 687 CAATGAACCTATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCTGATGTCCA 746
Qy 240 nTyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAlaVally 260
Db 747 ATACTTAGTAAATAAGCCTTCAGAAAGAGGGTCAGGAGATCATTTGTTGATGCTGTCAA 806
Qy 260 sIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIl 280
Db 807 AATTGAGCAGGAGTGTTTTAAACAGAGCCTTGCAGTGGCCTCATTTGGAATGAATGCAT 866
Qy 280 eLeuMetLysGlnTyrIleGluPheValAlaAspArgLeu-LeuValGluLeuGlyPheS 300
Db 867 TTTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAMCTTGGATCT 926
Qy 300 eLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyL 320
Db 927 CAAGAGKTTTTCAGGCAGAAAAATCCCTTTGATTTATGGAAM-AITTTCTTTAGAGGAA 985
Qy 320 yThrAsnPhePhe 324
Db 986 AAMA-AAATTTCTTT 998
```

RESULT 6

AY398973
LOCUS
DEFINITION
Homo sapiens HCM0069 gene, VIRUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
AY398973

AY398973.1 GI:39754962
GSS.
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
1 (bases 1 to 1056)
Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL
Science 302 (5652), 1960-1963 (2003)
PUBMED
14671302
REFERENCE
AUTHORS
2 (bases 1 to 1056)
Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE
Direct Submission
JOURNAL
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..1056
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1056
/locus_tag="HCM0069"
ORIGIN
Alignment Scores:
Pred. No.: 3,02e-182 Length: 1056
Score: 1585.00 Matches: 313
Percent Similarity: 89.17% Conservative: 0
Best Local Similarity: 89.17% Mismatches: 38
Query Match: 87.04% Indels: 0
DB: 9 Gaps: 0
US-10-698-228-1 (1-351) x AY398973 (1-1056)
Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 1 ATGGCGAGCCCGGAAGAGCGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Qy 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
Db 61 GACACCAACGAAGTGAATTAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheLysMetTyrLysGlnAlaGln 60
Db 121 CCGTTTTGTCATCTTTCCAAATCCCAATCCAGTACCTCGATATTTGGAAAATGTATAAACAGGCACA 180
Qy 61 AlaserPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrPheSerLys 80
Db 181 GCTTCTCTTCGACAGCAGAGAGTGCATTTATCAAAAGGATCTCCCTCAGTCGGAACAG 240
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 241 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGTAT 300
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 301 GGAATTCGTAATGAAAATTTGGTGAGCGCTTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160

```

Db 421 CTGATGACACTTACATCAGAGATCCAGAAAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 480
Qy 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540
Qy 181 SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
Db 541 NNNNNNNNNNGGAAAGAGTGGTGCCTTTGCTCTGTAGAGAGATTTCTCTCAGGA 600
Qy 201 SerPheAlaLalIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGTCTCTATTTCTGGCTAAAGAGAGGCTTATGCGCAGGACTCCTTTTTC 660
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTTTCAA 720
Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 721 TACTTAGTAATAAGCCTTCAGAAAGAGGTCAGGAGATCAATTGTTGATGCTGTCAA 780
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 781 ATTGACAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 840
Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 841 TTGATGAAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGGAACTTGGATTCTCA 900
Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 901 AAGGTTTTTTCAGGCAGAAAATCCTTTTGAATTTATGGAACAATTTCTTTAGAGGAAA 960
Qy 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 961 ACAAAATTTCTTTGAGAAACAGATTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
Qy 341 ThrAsnValPheThrLeuAspAlaAspPhe 351
Db 1021 ACAGATAACGCTCTTCACTTGGATGCGAGATTTT 1053

RESULT 7
LOCUS AY398975
DEFINITION Mus musculus HCM0069 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY398975
VERSION AY398975.1 GI:39754964
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1043)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1043)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering

```

```

them based on alignment.
FEATURES             Location/Qualifiers
     source            1..1043
     /organism="Mus musculus"
     /mol_type="genomic DNA"
     /db_xref="taxon:10090"
     gene              <1..>1043
     /locus_tag="HCM0069"

ORIGIN
Alignment Scores:
Pred. No.:          5,75e-168      Length:      1043
Score:              1468.00         Matches:    290
Percent Similarity: 86.61%         Conservative: 14
Best Local Similarity: 82.62%      Mismatches: 43
Query Match:        80.62%         Indels:     5
DB:                  9              Gaps:        1

US-10-698-228-1 (1-351) x AY398975 (1-1043)
Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 1 ATGGCGACCCGGAAGGCGCGAGCGCCAGCGGAGAGGGTGAGCAATGTGTTCG 60
Qy 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
Db 61 GAGACAGAAGAAATGTGTCAGGTCAAATGAAGAGCCACTCTTAAGAAAGAGTTC 120
Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysGlnAlaGln 60
Db 121 CGGTTGCGTCATCTTCCAAATCAGATCTCTGATATCTGGAGAAATGTACAGCAAGCAG 180
Qy 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
Db 181 GCATCTTCTCGACAGCAGAGAGGTTGACTTGTCAAGAGACTCCCTCCTCGAACAAG 240
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
Db 241 CTTAATCCGACGAGAGAGTATTTATCTCCACATCTTAGCCTTTTTCAGCCAGCGAT 300
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 301 GGAATTTGTAATGAAACTTGTGTCAGCGCTTTTAGTCAGAGGTCGACGCTTCAGAGGCT 360
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 361 CGCTGCTCTATGCGCTTTCAGATTCTCATCGAAGATGTTTCACTCAGAGATGTACAGTTA 420
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 421 CTATAGACACTTACATCAGAGATCCAAAGAAAGNNNNNNNNNNNNNNNNNNNNNNNNNN 480
Qy 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540
Qy 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db 541 NNNNNNNNNNGGAAAGAGTGGTGCCTTTGCTGCTGTGGAAGAGATTTCTCTCTGGA 600
Qy 201 SerPheAlaLalIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCATTTGCTGCATATTTCTGGCTAAAGAGAGGCGCTCATGCTGAGACTGACTTTTTC 660
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACATCATCAGCAGGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTTTCAA 720
Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 721 TACTTGTGTAACAAGCCTTCAGAAAGATAGAGTGGAGAAATCATTTGCTGATGCTGTCAA 780
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280

```



```

|||||
781 ATCGAGCA-----GAAGCCTTGCCTGTTGGCTCATTGGATGAATGTGTT 827
|||||
281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGlyPheSer 300
|||||
828 TTGATGAACAGTACATTGAGTTGTAGCTGACAGATTACTTGGAGAACTTGGATTCTCA 887
|||||
301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
|||||
888 AAGATTTTTCAGCGAGAAAATCCCTTGTATTTATGGAGAACATTTCAATTAGAGGGAAA 947
|||||
321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
|||||
948 ACAAAATTTCTTTGAGAAACGAGTTTCTGAGTATCAGCGATTTCGATCGTACGCCGAAACC 1007
|||||
341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
|||||
1008 ACAGATAATGTCTTACCTTGGATGCGAGATTTC 1040

CR603461 1571 bp mRNA linear HTC 21-JUL-2004
full-length cDNA clone CS0DH002YP23 of T cells (Jurkat cell line)
of Homo sapiens (human).
ACCESSION CR603461
VERSION CR603461.1 GI:50484268
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
2 (bases 1 to 1571)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1..1571
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DH002YP23"
/tissue_type="T cells (Jurkat cell line)"
/plasmid="pCMVSPORT_6"

ORIGIN
Alignment Scores:
Pred. No.: 8-86e-167 Length: 1571
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 3 Gaps: 1

US-10-698-228-1 (1-351) x CR603461 (1-1571)

RESULT 9
LOCUS CR625489 1573 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DA003YL09 of Neuroblastoma of Homo
sapiens (human).

```

```

QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
DB 204 GGAGCGGAACCTAAAGAGAGCTGCCCGCGCGTGGAGGATGAGCCGCTGCTGAGAGAAA 263
QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysG1 58
DB 264 CCCCCCGCGCTTGTCTATCTTCCCCTCGAGTACCATGATATCTGGCAGATGATATAAGAA 323
QY 58 nAlaGlnAlaSerPheThrAlaGluGluValAlaAspLeuSerLysLeuProHisTyr 78
DB 324 GGCAGAGGCTTCTCTTTGGACCGCCGAGGAGGTGGACCTCTCCAAGGACATTCAGCACTG 383
QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaA1 98
DB 384 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCATGTTCTTGGCTTCTTCTTGCAGC 443
QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
DB 444 AGCGATGGCATAGTAATGAAACCTTGGTGGAGGATTTAGCCAGAGTTTCAGATTAC 503
QY 118 oGluAlaArgCysPheThrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
DB 504 AGAAGCCCGCTGTTCTATGCTTCCAAATTTGCCATGGAACATACATCTGAAATGTA 563
QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
DB 564 TAGTCTCTTATTGACACTTACATATAAAGATCCCAAGAAAGGGAATTTCTTCTCAATGC 623
QY 158 alieGluThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrPheAlaAs 178
DB 624 CATGAAACGATGCCCTTGTGTCAAGAAGAGGAGACTGGCCCTTGGCTGGATGGGA 683
QY 178 pArgLysSerThrPheGlyGluArgValAlaAlaPheAlaAlaValGluGlyValPhePh 198
DB 684 CAAAGAGGCTACCTATGTGTAACGTGTGTAGCTTGTCTGCTGAGGAGGCAATTTCTTCT 743
QY 198 eSerGlySerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuTh 218
DB 744 TTCGGTCTTTTTCGGTCGATATTTCTGGCTCAAGAAACGAGGACTGTATGCTGGCTCAC 803
QY 218 rPheSerAsnGluIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
DB 804 ATTTTCTAATGAACCTTATTAGCAGAGATGAGGGTTTACACTGTGATTTTGTCTGCTGAT 863
QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
DB 864 GTTCAACACCTGGTACACAAACCTCGAGGAGAGTAAAGAGAAATAATATCAATGC 923
QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
DB 924 TGTTCGATAGAACAGGAGTTCTCTCACTGAGGCTTCTGCTGTAAGCTCATTTGGGATGAA 983
QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuG1 298
DB 984 TTGCACCTTAATGAAGCAATACATTTGAGTTGTGGCAGACAGACTTATGCTGGAACTGGG 1043
QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
DB 1044 TTTTAGCAAGGTTTTCAGATGAGAAACCCATTTGACTTTATGGAGATATTTCACTGGA 1103
QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
DB 1104 AGGAAAGACTTAACCTCTTTTGAAGAGAGATGAGGAGTATCAGAGGATGGGAGTGTATGTC 1163
QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1164 AAGTCCCAACAGAGAAATTTCTTTTACCTTGGATGCTGACTTC 1203

RESULT 9
LOCUS CR625489 1573 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DA003YL09 of Neuroblastoma of Homo
sapiens (human).

```



```

ACCESSION CR625489
VERSION CR625489.1 GI:50506296
KEYWORDS HTC; CNSLIT_CDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1573)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1573)
Genoscope.
REFERENCE Direct Submission
AUTHORS Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
TITLE BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
JOURNAL - Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
Location/Qualifiers
1..1573
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DA003YL09"
/tissue type="Neuroblastoma"
/plasmid="pCMVSPORT_6"
ORIGIN
Alignment Scores:
Pred. No.: 8,87e-167 Length: 1573
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 3 Gaps: 1
US-10-698-228-1 (1-351) x CR625489 (1-1573)
QY 4 ProGluArgProGluAlaGlyLeuAspGluGlnAspGluArgSerSerAspThrAs 23
Dy 155 CTGAGCGGAGCCCGCGTCTCGCCAGCAAGACCGGAGGAGGATCTCCAGGAGCCAC 214
QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuLysSe 38
Dy 215 GGAGCGGAAACCTAAGACGAGTCCCGCGGCGTGGAGGATGAGCCCTGCTGAGAGAAA 274
QY 38 rSerArgArgPheValIlePheProIleGlnTyProAspIleTrpLysMetTyLysG 58
Dy 275 CCCCAGCGCTTGTTCATCTTCCCATCGAGTACCATGATATCTGCAGATGATAGAA 334
QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
Dy 335 GGCAGAGGCTTCTTTTGGACCGCCGAGGAGGTGGACCTCTCCAGGACATTCAGCACTG 394
QY 78 pAsnLysLysLysAlaAspGluLysTyPheIleSerHisIleLeuAlaPheAlaAl 98
Dy 395 GGAATCCCTGGAACCGGAGAGATATTTATATCCCATGTTCTGGCTTTCTTGCAGC 454
QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Dy 455 AAGCGATGCGATGTAATGAAACCTTGTGGAGCATTTAGCCAGAGATTCAGATTAC 514
QY 118 oGluAlaArgCysPheTyrglyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Dy 515 AGAAGCCCGCTGTTCTATGGCTTCCAAATTCGATGGAACATACATCTCGAAATGTA 574

```

138 rSerLeuLeuIleAspThrTyrlleArgAspProIlyLysArgGluPheLeuPheAsnAl 158
575 TAGTCTTCTTATTGACACTTACATAAAGATGCCCAAGAAAGGGAATTTCTCTCAATGC 634
QY 158 aileGluThrMetProTyrvallLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
Dy 635 CATTGAACCGATGCTTGTCTCAAGAGAGGAGGAGCTGGGCTTGGCTGGATGGGGA 694
QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
Dy 695 CAAAGAGGCTACCTATGCTGAACGCTTGTAGCTTGTCTCAGTGAAGGCAATTTCTT 754
QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Dy 755 TTCGGTCTTCTTTCGTCGATATTCTGCTCAAGAAACGAGGACTGATGCTGGGCTCAC 814
QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Dy 815 ATTTCTTAATGAACCTTATTAGCAGATGAGGGTTTACACTGTGATTTTGTCTGCTGAT 874
QY 238 tPheGlnTyLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Dy 875 GTTCAACACCTGTCACAAACCATCGGAGGAGAGTAAAGAGAAATATTTATCAATGC 934
QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Dy 935 TGTTGGATAGAACAGGAGTTCTCTCACTGAGGCTTGTGCTGGAAGCTCATTTGGGATGA 994
QY 278 nCysIleLeuMetLysGlnTyrlleGluPheValAlaAspArgLeuLeuValGluLeuGl 298
Dy 995 TTGCACCTTAATGAAGCAATACATTTGTTGTGCGACAGACACTTATGTGGAACTGGG 1054
QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
Dy 1055 TTTTACGACAGTTCCTGAGTAGAGAACCCATTTGACTTTATGGAGATATTTTCACTGA 1114
QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrglnArgPheAlaValMetAl 338
Dy 1115 AGGAAGACTTAACCTCTTTCGAGAGAGAGTAGGCGAGTATCAGAGAGTGGGAGTGTATGC 1174
QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Dy 1175 AAGTCCCAACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1214

RESULT*10
CR608076 1582 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DK012YD23 of HeLa cells Cot 25-normalized
DEFINITION of Homo sapiens (human).
ACCESSION CR608076
VERSION CR608076.1 GI:50488883
KEYWORDS HTC; CNSLIT_CDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1582)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1582)
Genoscope.
REFERENCE Direct Submission
AUTHORS Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
TITLE BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
JOURNAL - Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a

```

division of Invitrogen.
FEATURES             Location/Qualifiers
     source           1..1582
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DA011YH14"
                     /issue="HeLa cells Cot 25-normalized"
                     /plasmid="pCMVSPORT_6"

ORIGIN
Alignment Scores:
Pred. No.:      8,95e-167      Length:      1582
Score:          1460.50        Matches:     277
Percent Similarity: 88.42%      Conservative: 36
Best Local Similarity: 78.25%    Mismatches:  35
Query Match:     80.20%        Indels:      6
DB:              3              Gaps:         1

US-10-698-228-1 (1-351) x CR608076 (1-1582)

QY      4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
Db      142 CTGAGCGGACCGCGTCTGCGCAGCAAGACCGGAGGAGATCTTCAGGAGGCCAC 201

QY      23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLys 38
Db      202 GGAGCGGAAACTAAAGCAGCTGCCCGCGGAGGAGTATGCGCTGAGAGAAA 261

QY      38 rSerArgArgPheValIlePheProIleGlnTyProAspIleTrpLysMetTyLysG 58
Db      262 CCCCCCGCGTTGTCTATCTCCCATCGATGATATCTGCAGATGTATAAGA 321

QY      58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHis 78
Db      322 GGCAGAGAGCTTCTTTTGGACCCCGGAGGAGGACCTCTCCAGAGCAATTCAGCACTG 381

QY      78 pAsnLysLeuLysAlaAspGluLysTyPheIleSerHisIleLeuAlaPhePheAla 98
Db      382 GGAATCCCTGAACCCGAGGAGAGATATTTATATCCATGTTCTGGCTTCTTCGAGC 441

QY      98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db      442 AAGCGATGCGCATAGTAATAATGAAACCTTGGTGGAGCGATTTAGCCAAAGATTCAGATTAC 501

QY      118 oGluAlaArgCysPheTyGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db      502 AGAAGCCCGCTGTTCTATGGCTTCCAAATTCGATGGAAACATACATCTCGAAATGTA 561

QY      138 rSerLeuLeuIleAspThrTyIleArgAspProLysLysLysArgGluPheLeuPheAsnAl 158
Db      562 TAGTCTTCTATTGACACTTACATATAAGATCCCAAGAAGGGAATTTCTCTTCATGC 621

QY      158 alleGluThrMetProTyTyValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
Db      622 CATTGAAACGATCCCTTGTGTCAAGAAGAAGGAGGAGCTGGCGCTTGGCGTGGATGGGA 681

QY      178 pArgLysSerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePh 198
Db      682 CAAAGAGGCTACCTATGTGTGAAGCTGTGTGAGCTTGTGCTGAGTGGAGGCAATTTCTT 741

QY      198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db      742 TTCGGTCTTTTGGCTGCGATATTCGGCTCAAGAACAGGAGGACTGATGCTGCGCTCAC 801

QY      218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db      802 ATTTCTAATGAATATTATAGCAGATGAGGGTTTACACTGTGATTTGCTTGGCTGAT 861

QY      238 tPheGlnTyLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db      862 GTTCAAAACACCTGGTACAAACCATCGGAGGAGAGATGAAGAGAAATATATTATCAATGC 921

```

```

QY      258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValClyLeuIleGlyMetAs 278
Db      922 TGTTCGATAGAACAGAGGTTCTCTACTGAGGCGCTTCCCTGTGAAGTCTCATTTGGATGAA 981

QY      278 nCysIleLeuMetLysGlnTyIleGluPheValAlaAspArgLeuLeuValGluLeuG 298
Db      982 TTGCATCTCTATGAGCAATACATTGATTTTGGCAGACAGACTTATCTGGAACCTGGG 1041

QY      298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG 318
Db      1042 TTTTAGCAGGTTTTCAGAGTAGAACCCATTGACTTTATGAGATAATTTCACTGCA 1101

QY      318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyGlnArgPheAlaValMetAl 338
Db      1102 AGGAAAGACTTAATCTTTTGGAGAGAGAGTAGGCGAGTATCAGAGGATGGAGTGTGTC 1161

QY      338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db      1162 AGTCCACACAGAGATTTCTTTTACCTTGGATGTGACTTC 1201

RESULT 11
CR602054      1588 bp      mRNA      linear      HTC 21-JUL-2004
LOCUS      full-length cDNA clone CS0DA011YH14 of Neuroblastoma of Homo
DEFINITION      sapiens (human) .
ACCESSION      CR602054
VERSION      CR602054.1 GI:50482861
KEYWORDS      HTC; CNSLT cDNA.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      1 (bases 1 to 1588)
AUTHORS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLES      Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL      Full-length cDNA libraries and normalization
REMARK      Unpublished
      Contact : Feng Liang Email : fliang@lifetech.com URL :
      http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
      Faraday Avenue
      Genoscope.
REFERENCE      2 (bases 1 to 1588)
AUTHORS      Direct Submission
TITLES      Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL      BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
      - Web : www.genoscope.cns.fr)
COMMENT      1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
      end enriched, double-strand cDNA was digested with Not I and cloned
      into the Not I and EcoK V sites of the pCMVSPORT 6 vector. Library
      was normalized. Library was constructed by Life Technologies, a
      division of Invitrogen.
FEATURES             Location/Qualifiers
     source           1..1588
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DA011YH14"
                     /tissue type="Neuroblastoma"
                     /plasmid="pCMVSPORT_6"

ORIGIN
Alignment Scores:
Pred. No.:      9e-167      Length:      1588
Score:          1460.50        Matches:     277
Percent Similarity: 88.42%      Conservative: 36
Best Local Similarity: 78.25%    Mismatches:  35
Query Match:     80.20%        Indels:      6
DB:              3              Gaps:         1

US-10-698-228-1 (1-351) x CR602054 (1-1588)

QY      4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
Db      142 CCTGAGCGGACCGCGCTCTCTGCGCAGCAAGACCGGAGGAGATCTTCAGGAGGCCAC 201

```



```

Qy 258 aVallysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
935 TGTTGGATAGAACAGAGTTCTCTACTGAGGCTTGGCTGTGAAGCTCATTTGGATGAA 994
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuG1 298
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
995 TTGCACTTAATGAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACTGGG 1054
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1055 TTTTACAGAGTTTTCAGACTAGAACCCATTGTGACTTTATGGAGATATTTTCACTGGA 1114
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 318 uGlyLysThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetal 338
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1115 AGAAAGACTAACTTCTTTGAGAAGAGAGTAGGCGAGTATCAGAGGATGGAGTGATGTC 1174
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1175 AAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1214
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 14
LOCUS CR590959 1605 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0D8013YD22 of Placenta of Homo sapiens
(human).
ACCESSION CR590959
VERSION CR590959.1 GI:50471766
KEYWORDS HTC; CDSIT_cDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1605)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
Genoscope.
2 (bases 1 to 1605)
Direct Submission
Genoscope.
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source Location/Qualifiers
1..1605
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D8013YD22"
/tissue_type="Placenta"
/plasmid="pCMVSPORT_6"
ORIGIN
Alignment Scores:
Pred. No.: 9,15e-167 Length: 1605
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 3 Gaps: 1

US-10-698-228-1 (1-351) x CR590959 (1-1605)
Qy 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

150 CCTGAGCGGACCGCGTCTCTGGCCAGCAAGCCGAGGAGGATCTTCCAGAGCCAC 209
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
210 GGAGCCGAAACTAAAGCAGCTGCCCGCGGTGGAGGATAGCCGCTGCTGAGAGAAA 269
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 38 rSerArgArgPheValIlePhePheProIleGlnTyrProAspIleTyrLysMetTyrLysG1 58
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
270 CCCCGCGCGTCTCTCATCTTCCCCTAGATCCCATGATATCTGGCAGATGTATAAGAA 329
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 58 nAlaGlnAlaserPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
330 GGCAGAGCTTCTTTTGGACCGCGAGAGGTGGACCTCTCCAGGACATTCAGACATG 389
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
390 GGAATCCTGAAACCCGAGGAGAGATATTTATATATCCCATGTTCTGGCTTTCTTTGCAGC 449
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 98 aSerAspGlyIleValAsnGluLeuValGluArgPheSerGlnGluValGlnValPr 118
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
450 AAGCGATGGCATAGTAAATGAAAACTTGGTGGAGCGATTAGCCCAAGAAAGTTTCAGATTAC 509
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyr 138
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
510 AGAAGCCGCGTGTCTTATGGCTTCCAAATTTGCCATGGAAACATACATTCGAAATGPA 569
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
570 TAGTCTTCTTATGACACTTACATTAAGATCCCAAGAAAGGGAATTTCTTCAATGC 629
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 158 aileGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTyrPheAlaAs 178
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
630 CATTGAAACGATGCTTGTGTCAAGAAAGAGCGAGACTGGCGCTTGCCTGGATTGGGGA 689
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 178 pArgLysSerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePhe 198
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
690 CAAAGAGGCTACCTTGTGTGAACGTTGTGTAGCCCTTTCGTCAGTGGAGGCAATTTCTT 749
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 198 eSerGlySerPheAlaAlaIlePheTyrPheLysLysArgGlyLeuMetProGlyLeuTh 218
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
750 TTCCGGTCTTTTTCGGTCGATATTTCTGGCTCAAGAAACGAGAGCTGATGCTTGGCTCTAC 809
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
810 ATTTCTTAATGAATTTATAGCAGAGATGAGGGTTTACACTGTGATTTTGTGCTGCTGAT 869
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
870 GTTCAAAACACTGGTACACAAACCATCGGAGGAGAGAGTAAGAGAAATAATATCAATGC 929
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
930 TGTTGGATAGAACAGAGTTCTCTACTGAGGCTTGCCTGTGAAGCTCATTTGGATGAA 989
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG1 298
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
990 TTGCACTTAATGAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACTGGG 1049
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1050 TTTTAGCAAGGTTTTCAGAGTAGAAGAACCCATTTTATGAGTATTTTCACTGGA 1109
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 318 uGlyLysThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetal 338
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1110 AGAAAGACTAACTTCTTTGAGAAGAGAGTAGGCGAGTATCAGAGGATGGAGATGATGTC 1169
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1170 AAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1209
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 15
LOCUS CR614990

```

CR614990 1612 bp mRNA linear HTC 21-JUL-2004

```

DEFINITION full-length cDNA clone CS0DA009YB23 of Neuroblastoma of Homo
              sapiens (human).
ACCESSION   CR614990
VERSION     CR614990.1 GI:50495797
KEYWORDS    HTC; CNSLT cDNA.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1612)
AUTHORS     Li W.B., Gruber C., Jessee J. and Polayes D.
TITLE       Full-length cDNA libraries and normalizations
JOURNAL     Unpublished
REMARK      Contact : Feng Liang Email : fliang@lifetech.com URL :
             http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
             Faraday Avenue
REFERENCE   2 (bases 1 to 1612)
AUTHORS     Genoscope.
TITLE       Direct Submission
JOURNAL     Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
             BP 191 91006 Evry cedex - FRANCE (E-mail : segref@genoscope.cns.fr
             - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
             end enriched, double-strand cDNA was digested with Not I and cloned
             into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
             was normalized. Library was constructed by Life Technologies, a
             division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1..1612
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DA009YB23"
                     /tissue_type="Neuroblastoma"
                     /plasmid="pCMVSPORT_6"
ORIGIN
Alignment Scores:
Pred. No.:      9,21e-167      Length:      1612
Score:          1460.50        Matches:     277
Percent Similarity: 88.42%      Conservative: 36
Best Local Similarity: 78.25%    Mismatches:  35
Query Match:     80.20%         Indels:      6
DB:              3             Gaps:        1

US-10-698-228-1 (1-351) x CR614990 (1-1612)

QY      4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
Db      152 CCTGAGCGGACCCCGGCTCTGGCCAGCAAGACCGCGGAGGAGATCTTCCAGGAGCCAC 211
QY      23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db      212 GGAGCGGAACAACTAAGACAGCTGCCCGCGTGGAGGATGAGCGGCTGCTGAGAGAAA 271
QY      38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysG1 58
Db      272 CCCCCCGCGCTTGTCTATCTTCCCATCGATGATATCTGGCAGATGTATAGAA 331
QY      58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
Db      332 GCGAGAGGCTCTCTTTTGGACCGCCGAGGAGGTGGACCTCTCCAAGACATTCAGCACTG 391
QY      78 pAnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
Db      392 GGAATCCCTGAACCCGAGGAGAGATATTTATATCCATCTTCTGCGCTTCTTTGCGAG 451
QY      98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerHisIleLeuAlaPhePheAl 118
Db      452 AGCGATGGCATAGTAATAAGAAACTTGGTGGAGCGATTAGCCAGAAAGTTCAGATTAC 511
QY      118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138

```

```

Db      512 AGAAGCCCGCTGTTTCTATGGCTTCCAAATGGCCATGGAAACATACATTCTGAAATGTA 571
QY      138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db      572 TAGTCTTCTTATTGACACTTACATAAAAGATCCCAAGAAAGGGAATTTCTCTCAATGC 631
QY      158 aIleGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
Db      632 CATTGAAACGATGCTTGTGTCAAGAAAGAGGAGACTGGGCGCTTGGCTGTGATTTGGGA 691
QY      178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePhe 198
Db      692 CAAGAGCGCTACCTATGCTGNAACGTGTGTAGCTTTGCTGAGTGAAGGCAATTTCTT 751
QY      198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db      752 TTCGGTTCCTTTTGGCTCGATATTCTGCTCAAGAAACGAGGACTGTGCTGCGCTGCCTCAC 811
QY      218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db      812 ATTTTCTAATGAACCTTATTAGCAGAGATGAGGGTTTACACTGTGATTTGCTTGCTGAT 871
QY      238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db      872 GTTCAAAACCTGTGTCACAAACCATCGAGGAGAGAGTAGAGAAATAATTTATCAATGC 931
QY      258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db      932 TGTTGGATAGAACAGGAGTTCTCTCACTGAGGCTTGGCTGTGAAAGCTCATTTGGGATGAA 991
QY      278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAlaAspArgLeuValGluLeuGl 298
Db      992 TTGCACCTCTAATGAAGCAATACTTGTGTCGACAGACACTTATGCTGGAACCTGGG 1051
QY      298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
Db      1052 TTTTAGCAAGTTTTCAGAGTAGAGAACCCATTTGACTTTATGGAGAATATTTCACTGGA 1111
QY      318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
Db      1112 AGGAAAGACTAACTTCTTTGAGAAAGAGATAGGCGAGTATCAGAGGATGGGAGTGATGTC 1171
QY      338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db      1172 AAGTCCAACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1211

```

Search completed: June 16, 2005, 14:11:19
Job time : 3873 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	1053	100.0	1053	6	AR454868	Sequence
2	1053	100.0	1053	6	BD064764	Novel pro
3	1053	100.0	1053	6	BD093077	Novel pro
4	1053	100.0	1056	6	CQ174252	Sequence
5	1053	100.0	1056	9	AB166671	Homo sapi
6	1053	100.0	4955	6	AR454869	Sequence
7	1053	100.0	4955	6	BD064765	Novel pro
8	1053	100.0	4955	6	BD093078	Novel pro
9	1053	100.0	4955	9	AB036063	Homo sapi
10	1051.4	99.8	1053	6	AR454877	Sequence
11	1051.4	99.8	1053	6	BD064773	Novel pro
12	1051.4	99.8	1053	6	BD093086	Novel pro
13	1051.4	99.8	1081	6	AR454870	Sequence
14	1051.4	99.8	1081	6	BD064766	Novel pro
15	1051.4	99.8	1081	6	BD093079	Novel pro
16	1049.8	99.7	1601	6	BD156916	Primer fo
17	1049.8	99.7	1601	6	AX877905	Sequence
18	1049.8	99.7	1601	9	AK001965	Homo sapi
19	895.2	85.0	4532	10	BC058103	Mus muscu

Db 241 CTTAAAGCAGATGAGAAGTACTTCATCTCTACATCTTAGCCTTTTTCAGCAGTGAT 300
QY 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Db 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCAGAGAAATGTTCACTCAGAGATGTACAGTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAAAATTCATCAGAGAAATGTTCACTCAGAGATGTACAGTTG 420
QY 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTTGAA 480
Db 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTTGAA 480
QY 481 ACCATGCCCTATCTTAAGAAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAAAA 540
Db 481 ACCATGCCCTATCTTAAGAAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAAAA 540
QY 541 TCTAATTTTGGGAAAAGAGTGGTGGCTTTGCTGTAGAAAGAGTTCCTTCAGGA 600
Db 541 TCTAATTTTGGGAAAAGAGTGGTGGCTTTGCTGTAGAAAGAGTTCCTTCAGGA 600
QY 601 TCTTTTCTGCTATATCTGGCTTAAGAAAAGAGTTCCTTCAGGA 660
Db 601 TCTTTTCTGCTATATCTGGCTTAAGAAAAGAGTTCCTTCAGGA 660
QY 661 AATGAACCTATCAGCAGATGAAGACATTCATCTGACATTTGCTTGGCTGATGTTCCAA 720
Db 661 AATGAACCTATCAGCAGATGAAGACATTCATCTGACATTTGCTTGGCTGATGTTCCAA 720
QY 721 TACTTAGTAATTAAGCTTCAGAAAAGAGGTGAGGATGATCTTGTGATGCTGTCAA 780
Db 721 TACTTAGTAATTAAGCTTCAGAAAAGAGGTGAGGATGATCTTGTGATGCTGTCAA 780
QY 781 ATTGACAGAGATTTTAAAGAGCTTCCAGAGTGGCTTCACTTGAATGAATTTGATTT 840
Db 781 ATTGACAGAGATTTTAAAGAGCTTCCAGAGTGGCTTCACTTGAATGAATTTGATTT 840
QY 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATCTTGGAACTTGGATTTCA 900
Db 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATCTTGGAACTTGGATTTCA 900
QY 901 AAGTTTTTTCAGCAGAAAATCCTTTTGAATTTTATGAAAACATTTCTTTAGAGGAAA 960
Db 901 AAGTTTTTTCAGCAGAAAATCCTTTTGAATTTTATGAAAACATTTCTTTAGAGGAAA 960
QY 961 ACAAATTTCTTTGAGAAAAGAGTTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAAC 1020
Db 961 ACAAATTTCTTTGAGAAAAGAGTTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAAC 1020
QY 1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053

RESULT 2
BD064764
LOCUS BD064764
DEFINITION Novel protein and DNA thereof.
ACCESSION BD064764
VERSION BD064764.1 GI:22610367
KEYWORDS JP 2001269184-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1053)
Nakamura, Y., Arakawa, H. and Tanaka, H.
Novel protein and DNA thereof
Patent: JP 2001269184-A 1 02-OCT-2001;
YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD
OS Homo sapiens (human)

PN JP 2001269184-A/1
PD 02-OCT-2001
PI 27-JUN-2000 JP 2000192401
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N15/09, A61K31/711, A61K38/43, A61K45/00, A61K48/00, PC
A61P35/00.
PC A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC
, C12N5/10, C12N9/02,
PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53// (C12N15/09,
C12R1:91),
PC (C12N15/09, C12R1:19), (C12N9/02, C12R1:91), (C12P21/02, C12R1:91),
PC (C12N15/00, A61K37/02, A61K37/48, C12N5/00,
C12N15/00, C12R1:91), (C12N15/00, C12R1:19)
CC Novel protein and DNA thereof
FH Key Location/Qualifiers
FT source 1. 1053
/organism="Homo sapiens (human)".
FEATURES
source
1. 1053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 1053; DB 6; Length 1053;
Best Local Similarity 100.0%; Pred. No. 1.2e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGCACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGCGCACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 121 CGGTTTGTATCTTTCCAAATCCAGTACCTGATATTTGGAAATATGATAAAGCAGCAG 180
Db 121 CGGTTTGTATCTTTCCAAATCCAGTACCTGATATTTGGAAATATGATAAAGCAGCAG 180
QY 181 GCTTCTCTTGGACAGCAGAGAGGTGACATTAAGAGGATCTCCCTCAGTGGACAG 240
Db 181 GCTTCTCTTGGACAGCAGAGAGGTGACATTAAGAGGATCTCCCTCAGTGGACAG 240
QY 241 CTTAAGCAGATGAGAAGTACTTCTCTCACATCTTAGCCTTTTTCAGCAGCAGTAT 300
Db 241 CTTAAGCAGATGAGAAGTACTTCTCTCACATCTTAGCCTTTTTCAGCAGCAGTAT 300
QY 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Db 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCAGAGATTTTCACTCAGAGATGTACAGTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAAAATTCATCAGAGATTTTCACTCAGAGATGTACAGTTG 420
QY 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTTGAA 480
Db 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAGAGTGGCCCTTGGATGGATAGCAGATAGAAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAGAGTGGCCCTTGGATGGATAGCAGATAGAAAA 540
QY 541 TCTAATTTTGGGAAAAGAGTGGTGGCTTTGCTGCTGTAGAAAGAGTTCCTTCAGGA 600
Db 541 TCTAATTTTGGGAAAAGAGTGGTGGCTTTGCTGCTGTAGAAAGAGTTCCTTCAGGA 600
QY 601 TCTTTTCTGCTATATCTGGCTTAAGAAAAGAGTTCCTTCAGGA 660
Db 601 TCTTTTCTGCTATATCTGGCTTAAGAAAAGAGTTCCTTCAGGA 660
QY 661 AATGAACCTATCAGCAGATGAAGACATTCATCTGACATTTGCTTGGCTGATGTTCCAA 720

VERSION	CQ714252.1	GI:42275109	
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.		
AUTHORS	Kits, such as nucleic acid arrays, comprising a majority of		
TITLE	humanexons or transcripts, for detecting expression and other uses thereof		
JOURNAL	Patent: WO 02068579-A 186 06-SEP-2002;		
PE Corporation (NY) (US)			
FEATURES	Location/Qualifiers		
source	1..1056		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN	Query Match 100.0%; Score 1053; DB 6; Length 1056;		
	Best Local Similarity 100.0%; Pred. No. 1.2e-255;		
	Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	ATGGCGACCCGGAAAGCGGCGGCTGGATCAGATCAGATCATCTTCA	60
Db	1	ATGGCGACCCGGAAAGCGGCGGCTGGATCAGATCAGATCATCTTCA	60
QY	61	GACACCAAGAAAGTCAATTAAGTCAATGAGAGCCACTCTTAAGAAAGAGTTCTCGC	120
Db	61	GACACCAAGAAAGTCAATTAAGTCAATGAGAGCCACTCTTAAGAAAGAGTTCTCGC	120
QY	121	CGTTTGTTCATCTTCCATCCAGTACCCTGATATTTGGAAAAATGATAAACAAGCACAG	180
Db	121	CGTTTGTTCATCTTCCATCCAGTACCCTGATATTTGGAAAAATGATAAACAAGCACAG	180
QY	181	GCTTCCTCTCGACAGCAGAAAGGTCGACTTATCAAAAGATCTCCCTCACTGGAACAAG	240
Db	181	GCTTCCTCTCGACAGCAGAAAGGTCGACTTATCAAAAGATCTCCCTCACTGGAACAAG	240
QY	241	CTTAAGCAGATGAGAAGTACTTCATCTCTCACATCTTAGCCCTTTTTCGAGCCAGTGAT	300
Db	241	CTTAAGCAGATGAGAAGTACTTCATCTCTCACATCTTAGCCCTTTTTCGAGCCAGTGAT	300
QY	301	GGAAATGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT	360
Db	301	GGAAATGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT	360
QY	361	CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAATGTTTCACCTCAGAGATGTACAGTTTG	420
Db	361	CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAATGTTTCACCTCAGAGATGTACAGTTTG	420
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTATGCAATTGAA	480
Db	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTATGCAATTGAA	480
QY	481	ACCATGCCCTATGTTAAGAAAAGCAGATGCGCCTTCGGATGATAGCAGATAGAA	540
Db	481	ACCATGCCCTATGTTAAGAAAAGCAGATGCGCCTTCGGATGATAGCAGATAGAA	540
QY	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGTTTCTCTCAGGA	600
Db	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGTTTCTCTCAGGA	600
QY	601	TCTTTTGTGCTATATCTCGCTAAAGAGAGGTCCTTATGCAAGCACTCACTTTTTC	660
Db	601	TCTTTTGTGCTATATCTCGCTAAAGAGAGGTCCTTATGCAAGCACTCACTTTTTC	660
QY	661	AATGAATCATCAGCAGAGATGAGCACTTCACTGTGACTTTGCTGCTGATGTTCCAA	720
Db	661	AATGAATCATCAGCAGAGATGAGCACTTCACTGTGACTTTGCTGCTGATGTTCCAA	720
QY	721	TACTTAGTAAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA	780

Db	721	TACTTAGTAAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA	780
QY	781	ATTGACGACGAGTGTATTTAAACAGAAAGCCCTGCGAGTTGGCTCATTTGGAATGAATTGCATT	840
Db	781	ATTGACGACGAGTGTATTTAAACAGAAAGCCCTGCGAGTTGGCTCATTTGGAATGAATTGCATT	840
QY	841	TTGATGAAAACAGTACATTTGAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTCTCA	900
Db	841	TTGATGAAAACAGTACATTTGAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTCTCA	900
QY	901	AAGTTTTCAGGCAGAAAATCCTTTTGAATTTATGAAAACATTTCTTTAGAGGAAAA	960
Db	901	AAGTTTTCAGGCAGAAAATCCTTTTGAATTTATGAAAACATTTCTTTAGAGGAAAA	960
QY	961	ACAAATTTCTTTGAGAAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC	1020
Db	961	ACAAATTTCTTTGAGAAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC	1020
QY	1021	ACAGATAACGCTCTTCACTTGGATGCAGATTTT 1053	
Db	1021	ACAGATAACGCTCTTCACTTGGATGCAGATTTT 1053	

RESULT 5

AB166671

LOCUS

DEFINITION

1056 bp mRNA linear PRI 09-MAR-2004

Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase

small subunit 2, complete cds.

ACCESSION

AB166671

VERSION

AB166671.1

GI:45259568

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 Ugai, H. and Yokoyama, K.K.

AUTHORS

Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase

TITLE

small subunit 2, complete cds

JOURNAL

Published Only in Database (2004)

REFERENCE

2 (bases 1 to 1056)

AUTHORS

Ugai, H. and Yokoyama, K.K.

TITLE

Direct Submission

JOURNAL

Submitted (05-MAR-2004) Hideyo Ugai, RIKEN Bioresource Center, Gene Engineering Division; 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan (E-mail:ugai@rtc.riken.go.jp, Tel:81-29-836-3612, Fax:81-29-836-9120)

FEATURES

Location/Qualifiers

1..1056

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="p53R2.C19"

/cell_line="Saos-2"

/cell_type="Osteogenic sarcoma"

gene

1..1056

/gene="p53R2"

CDS

1..1056

/codon_start=1

/product="p53-inducible ribonucleotide reductase small subunit 2"

/protein_id="BAD12267.1"

/db_xref="GI:45259568"

/translation="MGDPPEAAGLDQDERSSDTNEIKNSEPLLRKSRFRVI
FNTQPDIDWMYKQAQAFWAEVDLSKDLPHWNKLKADKPYFISHLAFPAASDGI
VNEVLVERFQEQVPEARCFYGFQIILINHVSEMYSLLDITYIRDPKRRFLFNAIE
TWPYKKADMALRWLADKSTGERVVAFAAVEGVFSGSFAAI FWLKRGFLMPGLT
FSNELLSRDEGLHCDFACLMPOILVINKPSEERREIIVADVAKIEQELTEALPVGLIG
ANMCLMKQYIEFVAERLLVSLGFSKVFQENPFDMENI SLEGKTNF FEKRVSEYQRF
AVNAETTDNVFTLDAF"

ORIGIN

Query Match 100.0%; Score 1053; DB 9; Length 1056;
 Best Local Similarity 100.0%; Pred. No. 1.2e-255;
 Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
 DB 1 ATGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA 60

QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGTCTCGC 120
 DB 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGTCTCGC 120

QY 121 CGTTTGTCTATCTTTCAATCCAGTACCTGTATATTGGAAATGTATAAACAGGCACAG 180
 DB 121 CGTTTGTCTATCTTTCAATCCAGTACCTGTATATTGGAAATGTATAAACAGGCACAG 180

QY 181 GCTTCTCTTGGACAGACGAGGTGCACTTATCAAGAGTCTCCCTCACTGGAACAAG 240
 DB 181 GCTTCTCTTGGACAGACGAGGTGCACTTATCAAGAGTCTCCCTCACTGGAACAAG 240

QY 241 CTTAAAGCAGATGAGAAAGTACTTCTCTCACAATCTTAGCCCTTTTGGCAGCCAGTAT 300
 DB 241 CTTAAAGCAGATGAGAAAGTACTTCTCTCACAATCTTAGCCCTTTTGGCAGCCAGTAT 300

QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTTCAGAGGCT 360
 DB 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTTCAGAGGCT 360

QY 361 CGTGTTTCTATGCTTTCAAAATCTCATCGAGAAATGTTTCACTCAGAGATGTACAGTTG 420
 DB 361 CGTGTTTCTATGCTTTCAAAATCTCATCGAGAAATGTTTCACTCAGAGATGTACAGTTG 420

QY 421 CTGATAGACATTTACATCAGAGATCCNAGAAAGGGAATTTTATTTAATGCAATGAA 480
 DB 421 CTGATAGACATTTACATCAGAGATCCNAGAAAGGGAATTTTATTTAATGCAATGAA 480

QY 481 ACCATGCCCTATGTTAAGAAAGAACAGATTTGGCCCTTGGCATGGAATAGAGAA 540
 DB 481 ACCATGCCCTATGTTAAGAAAGAACAGATTTGGCCCTTGGCATGGAATAGAGAA 540

QY 541 TCTA TTTTGGGAAAGAGTGTGGCTTTGCTGTGTAGAAGAGTTTCTTCTCAGGA 600
 DB 541 TCTA TTTTGGGAAAGAGTGTGGCTTTGCTGTGTAGAAGAGTTTCTTCTCAGGA 600

QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGACTCACTTTTCC 660
 DB 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGACTCACTTTTCC 660

QY 661 AATGAATCTCATCAGCAGATGAAGACTTTCACTGTGACTTTTCTGCTGCTGATTTCCAA 720
 DB 661 AATGAATCTCATCAGCAGATGAAGACTTTCACTGTGACTTTTCTGCTGCTGATTTCCAA 720

QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTGTGCTGTCAAA 780
 DB 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTGTGCTGTCAAA 780

QY 781 ATTGACGAGAGTTTAAACAGAGCTTGGCAGAGTGGCCCTCATTTGGAATGGAATTGCA 840
 DB 781 ATTGACGAGAGTTTAAACAGAGCTTGGCAGAGTGGCCCTCATTTGGAATGGAATTGCA 840

QY 841 TTGATCAAAACGATACATTCAGTTTGTAGCTGACAGATTAAGTGTGGAACCTTGAATCTCA 900
 DB 841 TTGATCAAAACGATACATTCAGTTTGTAGCTGACAGATTAAGTGTGGAACCTTGAATCTCA 900

QY 901 AAGGTTTTTTCAGGAGAAATCTTTTGTATTTTATGAAATCATTTCTTAGAAGGAAA 960
 DB 901 AAGGTTTTTTCAGGAGAAATCTTTTGTATTTTATGAAATCATTTCTTAGAAGGAAA 960

QY 961 ACAAAATTTCTTTGAGAAAACAGTTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1020
 DB 961 ACAAAATTTCTTTGAGAAAACAGTTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1020

QY 1021 ACAGATAACGCTCTTCACCTTGGATGCGAGATTTT 1053
 DB 1021 ACAGATAACGCTCTTCACCTTGGATGCGAGATTTT 1053

RESULT 6
 AR454869
 LOCUS AR454869 4955 bp DNA linear PAT 20-FEB-2004
 DEFINITION Sequence 3 from patent US 6682917.
 ACCESSION AR454869
 VERSION AR454869.1 GI:42688824
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 4955)
 AUTHORS Nakamura, Y., Arakawa, H. and Tanaka, H.
 TITLE Protein having a ribonucleotide Reductase activity and a DNA thereof
 JOURNAL Patent: US 6682917-A 3 27-JAN-2004;
 FEATURES Location/Qualifiers
 source 1..4955
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN
 Query Match 100.0%; Score 1053; DB 6; Length 4955;
 Best Local Similarity 100.0%; Pred. No. 1.1e-255;
 Matches 1053; Conservative 0; Mismatches 0; Indels - 0; Gaps 0;

QY 1 ATGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
 DB 245 ATGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 304

QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAAGAAAGTCTCGC 120
 DB 305 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAAGAAAGTCTCGC 364

QY 121 CGTTTGTCTATCTTTCAATCCAGTACCTGTATATTGGAAATGTATAAACAGGCACAG 180
 DB 365 CGTTTGTCTATCTTTCAATCCAGTACCTGTATATTGGAAATGTATAAACAGGCACAG 424

QY 181 GCTTCTCTTGGACAGACGAGGTGCACTTATCAAGAGTCTCCCTCACTGGAACAAG 240
 DB 425 GCTTCTCTTGGACAGACGAGGTGCACTTATCAAGAGTCTCCCTCACTGGAACAAG 484

QY 241 CTTAAAGCAGATGAGAAAGTACTTCTCTCACAATCTTAGCCCTTTTGGCAGCCAGTAT 300
 DB 485 CTTAAAGCAGATGAGAAAGTACTTCTCTCACAATCTTAGCCCTTTTGGCAGCCAGTAT 544

QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTTCAGAGGCT 360
 DB 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTTCAGAGGCT 604

QY 361 CGCTGTTTCTATGCTTTCAAAATCTCATCGAGAAATGTCTCACTCAGAGATGTACAGTTG 420
 DB 605 CGCTGTTTCTATGCTTTCAAAATCTCATCGAGAAATGTCTCACTCAGAGATGTACAGTTG 664

QY 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 480
 DB 665 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 724

QY 481 ACCATGCCCTATGTTAAGAAAGAACAGATTTGGCCCTTGGCATGGAATAGAGAA 540
 DB 725 ACCATGCCCTATGTTAAGAAAGAACAGATTTGGCCCTTGGCATGGAATAGAGAA 784

QY 541 TCTA TTTTGGGAAAGAGTGTGGCTTTGCTGTGTAGAAGAGTTTCTTCTCAGGA 600
 DB 785 TCTA TTTTGGGAAAGAGTGTGGCTTTGCTGTGTAGAAGAGTTTCTTCTCAGGA 844

QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC 660
 DB 845 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC 904

```
QY 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTCCCTGATGTTCCAA 720
|
|
|
Db 905 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
|
|
|
QY 721 TACTTAGTAAATAAGCCTTCAGAGAGAAAGGTCAGGAGAGATCAATGTTGATGCTGTCAAA 780
|
|
|
Db 965 TACTTAGTAAATAAGCCTTCAGAGAGAAAGGTCAGGAGAGATCAATGTTGATGCTGTCAAA 1024
|
|
|
QY 781 ATTGAGCAGAGATTTTAAACAGAGCCTTCCAGTGTGCTCATTTGGAATGAATTCATTT 840
|
|
|
Db 1025 ATTGAGCAGAGATTTTAAACAGAGCCTTCCAGTGTGCTCATTTGGAATGAATTCATTT 1084
|
|
|
QY 841 TTGATGAACAGTACATGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTCA 900
|
|
|
Db 1085 TTGATGAACAGTACATGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTCA 1144
|
|
|
QY 901 AAGGTTTTTTCAGCAGAGAAATCCTTTGATTTTATGGAATGATTTTCTTTAGAGGAAA 960
|
|
|
Db 1145 AAGGTTTTTTCAGCAGAGAAATCCTTTGATTTTATGGAATGATTTTCTTTAGAGGAAA 1204
|
|
|
QY 961 ACAAAATTTCTTTCAGCAGAGAAATCCTTTGATTTTATGGAATGATTTTCTTTAGAGGAAA 1020
|
|
|
Db 1205 ACAAAATTTCTTTCAGCAGAGAAATCCTTTGATTTTATGGAATGATTTTCTTTAGAGGAAA 1264
|
|
|
QY 1021 ACAGATAACGTCCTTCACTTGGATGCGAGATTTT 1053
|
|
|
Db 1265 ACAGATAACGTCCTTCACTTGGATGCGAGATTTT 1297
|
|
|

RESULT 7
BD064765
LOCUS Novel protein and DNA thereof.
DEFINITION Novel protein and DNA thereof.
ACCESSION BD064765
VERSION BD064765.1 GI:22610368
KEYWORDS JP 2001269184-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE Novel protein and DNA thereof
JOURNAL Patent: JP 2001269184-A 2 02-OCT-2001;
YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD
COMMENT
OS Homo sapiens (human)
PN JP 2001269184-A/2
PD 02-OCT-2001
PF 27-JUN-2000 JP 2000192401
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N15/09,A61K31/711,A61K38/00,A61K38/43,A61K45/00,A61K48/00, PC
A61P35/00,
PC A61P43/00,C07K16/32,C07K16/40,C12N1/15,C12N1/19,C12N1/21 PC
C12N5/10,C12N9/02,
PC C12P21/02,C12P21/08,G01N33/15,G01N33/50,G01N33/53///C12N15/09,
PC C12R1.91)
PC C12N15/09,C12R1.19),(C12N9/02,C12R1.91),(C12P21/02,C12R1.91),
PC (C12P21/08,C12R1.91),C12N15/00,A61K37/02,A61K37/48,C12N5/00,
PC C12N15/00,C12R1.91),(C12N15/00,C12R1.19)
CC Novel protein and DNA thereof
FH Key Location/Qualifiers
FT source 1.4955
FT Location/Qualifiers
source 1.4955
/organism='Homo sapiens (human)'.
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 1053; DB 6; Length 4955;
Best Local Similarity 100.0%; Pred. No. 1.1e-255;
```

```
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGCAACCGGAAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
|
|
|
Db 245 ATGGCGCAACCGGAAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
|
|
|
QY 61 GACACCAACCAAGTGAATAAAGTCAANTGAAGGCCACTCCTTAAGAAAGGTTCTCGC 120
|
|
|
Db 305 GACACCAACCAAGTGAATAAAGTCAANTGAAGGCCACTCCTTAAGAAAGGTTCTCGC 364
|
|
|
QY 121 CGGTTTCTCATCTTTCATATCCAGTACCTCATATTTGGAAAAATGTATAAACAGGCACAG 180
|
|
|
Db 365 CGGTTTCTCATCTTTCATATCCAGTACCTCATATTTGGAAAAATGTATAAACAGGCACAG 424
|
|
|
QY 181 GCTTCTCTTGGACAGCAGAGAGGTCACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 240
|
|
|
Db 425 GCTTCTCTTGGACAGCAGAGAGAGGTCACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 484
|
|
|
QY 241 CTTAAAGCAGATGAGAAAGTACTTCTATCTCTACATCTTAGCCTTTTTCAGGCCAGTGAT 300
|
|
|
Db 485 CTTAAAGCAGATGAGAAAGTACTTCTATCTCTACATCTTAGCCTTTTTCAGGCCAGTGAT 544
|
|
|
QY 301 GGAATTCGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGACGTTCCAGAGGT 360
|
|
|
Db 545 GGAATTCGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGACGTTCCAGAGGT 604
|
|
|
QY 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAGAAATGTTCACTCAGAGATGTACAGTTG 420
|
|
|
Db 605 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAGAAATGTTCACTCAGAGATGTACAGTTG 664
|
|
|
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTGAA 480
|
|
|
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTGAA 724
|
|
|
QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCTTTCCGATGATGATAGCAGATAGAAA 540
|
|
|
Db 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCTTTCCGATGATGATAGCAGATAGAAA 784
|
|
|
QY 541 TCTACTTTTGGGAAAAAGAGTGGTGGCTTTCTCTGTAGAGAGGAGTTTCTTCTCAGGA 600
|
|
|
Db 785 TCTACTTTTGGGAAAAAGAGTGGTGGCTTTCTCTGTAGAGAGGAGTTTCTTCTCAGGA 844
|
|
|
QY 601 TCTTTTCTGCTATATTTCTGTGCTAAAGAGAGGTCTTATGTCAGGAGCTCACTTTTTC 660
|
|
|
Db 845 TCTTTTCTGCTATATTTCTGTGCTAAAGAGAGGTCTTATGTCAGGAGCTCACTTTTTC 904
|
|
|
QY 561 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
|
|
|
Db 905 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
|
|
|
QY 721 TACTTAGTAAATAAGCCTTCAGAGAGAAAGGTCAGGAGAGATCAATGTTGATGCTGTCAAA 780
|
|
|
Db 965 TACTTAGTAAATAAGCCTTCAGAGAGAAAGGTCAGGAGAGATCAATGTTGATGCTGTCAAA 1024
|
|
|
QY 781 ATTGAGCAGAGATTTTAAACAGAGCCTTCCAGTGTGCTCATTTGGAATGAATTCATTT 840
|
|
|
Db 1025 ATTGAGCAGAGATTTTAAACAGAGCCTTCCAGTGTGCTCATTTGGAATGAATTCATTT 1084
|
|
|
QY 841 TTGATGAACAGTACATGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTCA 900
|
|
|
Db 1085 TTGATGAACAGTACATGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTCA 1144
|
|
|
QY 901 AAGGTTTTTTCAGGCGAGAGAAATCCTTTGATTTTATGGAATGATTTTCTTTAGAGGAAA 960
|
|
|
Db 1145 AAGGTTTTTTCAGGCGAGAGAAATCCTTTGATTTTATGGAATGATTTTCTTTAGAGGAAA 1204
|
|
|
QY 961 ACAAAATTTCTTTCAGAGAAACGAGTTTTCAGAGTATCAGCGCTTTTTCAGAGTATGCGAGAAACC 1020
|
|
|
Db 1205 ACAAAATTTCTTTCAGAGAAACGAGTTTTCAGAGTATCAGCGCTTTTTCAGAGTATGCGAGAAACC 1264
|
|
|
QY 1021 ACAGATAACGTCCTTCACTTGGATGCGAGATTTT 1053
|
|
|
Db 1265 ACAGATAACGTCCTTCACTTGGATGCGAGATTTT 1297
|
|
|
```

```

RESULT 8
BD093078
LOCUS
DEFINITION Novel protein and its DNA.
ACCESSION BD093078
VERSION BD093078.1 GI:22638666
KEYWORDS WO 0100799-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 4955)
AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE Novel protein and its DNA
JOURNAL Patent: WO 0100799-A 2 04-JAN-2001;
TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,
HIROSHI TANAKA
COMMENT
OS Homo sapiens (human)
PN WO 0100799-A/2
PD 04-JAN-2001
PF 27-JUN-2000 WO 2000JP004189
PR 28-JUN-1999 JP 99P 181131,06-JUL-1999 JP 99P 192391 PR
21-JAN-2000 JP 00P 017770
PI YUSUKE NAKAMURA,HIROFUMI ARAKAWA,HIROSHI TANAKA PC
C12N9/04,C12N15/53,C12N1/15,C12N1/19,C12N5/21,C12N5/10 PC
.C12P21/02,A61K38/44,
PC A61K48/00,C07K16/40,G01N33/53,C12Q1/26
CC
FH Key Location/Qualifiers.
FEATURES
source
1..4955
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 1053; DB 6; Length 4955;
Best Local Similarity 100.0%; Pred. No. 1.1e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGACCGGAAAGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 245 ATGGCGACCGGAAAGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
QY 61 GACACCAACGAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 305 GACACCAACGAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
QY 121 CGGTTTGTCACTTTCCAAATCCAGTACCTGATATTGGAAATGTATAACAGGCACAG 180
DB 365 CGGTTTGTCACTTTCCAAATCCAGTACCTGATATTGGAAATGTATAACAGGCACAG 424
QY 181 GCTTCTCTTGGACAGCAGAAAGAGTGCATTTAAAGAGATCTCCCTCACTGGAAACAAG 240
DB 425 GCTTCTCTTGGACAGCAGAAAGAGTGCATTTAAAGAGATCTCCCTCACTGGAAACAAG 484
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTTTAGCCTTTTTCGAGCCAGTGAT 300
DB 485 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTTTAGCCTTTTTCGAGCCAGTGAT 544
QY 301 GGAATGTAAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGCT 360
DB 545 GGAATGTAAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGCT 604
QY 361 CGCTGTTTCTATGCGCTTTCAAAATTTCTCATCGAGAACTGTTCACTCAGAGATGTACAGTTTG 420
DB 605 CGCTGTTTCTATGCGCTTTCAAAATTTCTCATCGAGAACTGTTCACTCAGAGATGTACAGTTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 480
DB 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 724

```

```

QY 481 ACCATGCCCTATGTTAAGAAAAAACAGATTTGGGCTTGGGATGATAGCAGATAGAAAA 540
DB 725 ACCATGCCCTATGTTAAGAAAAAACAGATTTGGGCTTGGGATGATAGCAGATAGAAAA 784
QY 541 TCTACTTTTGGGAAAGAGTGTGCTTGTCTGTAGAGAGGAGTTTCTTCTCAGGA 600
DB 785 TCTACTTTTGGGAAAGAGTGTGCTTGTCTGTAGAGAGGAGTTTCTTCTCAGGA 844
QY 601 TCTTTTGTCTATATTCTGGCTAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
DB 845 TCTTTTGTCTATATTCTGGCTAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 904
QY 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 720
DB 905 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 964
QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTTCAGGAGATCATTTGTGATGCTGTCAA 780
DB 965 TACTTAGTAAATAAGCCTTCAGAAAGAGGTTCAGGAGATCATTTGTGATGCTGTCAA 1024
QY 781 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTCGATT 840
DB 1025 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTCGATT 1084
QY 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTTGTGGAACTTGGATTCTCA 900
DB 1085 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTTGTGGAACTTGGATTCTCA 1144
QY 901 AAGGTTTTTCAGGCGAGAAATCCTTTGATTTATGAAACATTTCTTTAGAGAAAA 960
DB 1145 AAGGTTTTTCAGGCGAGAAATCCTTTGATTTATGAAACATTTCTTTAGAGAAAA 1204
QY 961 ACAATTTCTTTGAGAACGAGTTTCAGAGATATCAGCGTTTTTGCAGTTATGCGAGAAACC 1020
DB 1205 ACAATTTCTTTGAGAACGAGTTTCAGAGATATCAGCGTTTTTGCAGTTATGCGAGAAACC 1264
QY 1021 ACAGATAACGCTTTCACCTTGGATGCGAGATTTT 1053
DB 1265 ACAGATAACGCTTTCACCTTGGATGCGAGATTTT 1297

RESULT 9
AB036063
LOCUS
DEFINITION Homo sapiens p53R2 mRNA for ribonucleotide reductase, complete cds.
ACCESSION AB036063
VERSION AB036063.1 GI:7229085
KEYWORDS p53R2; ribonucleotide reductase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (sites)
AUTHORS Tanaka,H., Arakawa,H., Yamaguchi,T., Shiraishi,K., Fukuda,S.,
Matsui,K., Takei,Y. and Nakamura,Y.
TITLE A ribonucleotide reductase gene involved in a p53-dependent
cell-cycle checkpoint for DNA damage
JOURNAL Nature 404 (6773), 42-49 (2000)
MEDLINE 20179179
PUBMED 10716435
REFERENCE
2 (bases 1 to 4955)
AUTHORS Tanaka,H., Arakawa,H. and Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (18-DEC-1999) Yusuke Nakamura, University of Tokyo,
Institute of Medical Science, Human Genome Center, Laboratory of
Molecular Medicine; 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639,
Japan (E-mail:yusuke@ims.u-tokyo.ac.jp, Tel:+81-3-5449-5372,
Fax:+81-3-5449-5433)
FEATURES
Location/Qualifiers
source
1..4955
/organism="Homo sapiens"
/mol_type="mRNA"

```



```
/db_xref="taxon:9606"
/chromosome="8"
/map="8q23.1"
1..4955
/gene="p53R2"
245..1300
/gene="p53R2"
/codon_start=1
/product="ribonucleotide reductase"
/protein_id="BA92434.1"
/db_xref="GI:7229086"
/translation="MGDPERPEAAGLDODERSSDTNESPFIKSNBEPILRKSSRFVI
FPIQPDIMWTKQASFTAEEDVLSKDLPHWNKLKADEKFIHILAPFAASDGI
VNNELVERSQVQPEARCFYQFQILINVSHEMYSLIDITYIRDPKKRELFNLAIE
TMPYVKKADWALRIADRKSFTGERVAFVAVEGFFSGSFAAI FPKLKKRELFNLAIE
FNNELI SRDGEHLCDPACILMFOVLNKPSEERVREIIVDAVKIEQEFLEALPVGILG
MNCILMKOYIEFVADRLIVELGFSKVFQAEINPDMENISLEKTNPFPEKRVSEYQRF
AVMAETDNTDVLDDPF"
4933
/polyA_site
/gene="p53R2"
/note="22 a nucleotides"

ORIGIN
Query Match 100.0%; Score 1053; DB 9; Length 4955;
Best Local Similarity 100.0%; Pred. No. 1.1e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGACCCGGAAGCGGAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
DB 245 ATGGCGACCCGGAAGCGGAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 304
QY 61 GACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 120
DB 305 GACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 364
QY 121 CGGTTTGTCATCTTTCATATCCAGTACCTCTGATATTTGGAAATGATATAACAGGCACAG 180
DB 365 CGGTTTGTCATCTTTCATATCCAGTACCTCTGATATTTGGAAATGATATAACAGGCACAG 424
QY 181 GCTTCTCTTGGACAGCAGAGAGTGCAGTATCAAGAGTCTCCTCACTGCAACAG 240
DB 425 GCTTCTCTTGGACAGCAGAGAGTGCAGTATCAAGAGTCTCCTCACTGCAACAG 484
QY 241 CTTAAAGCAGATCAGAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 300
DB 485 CTTAAAGCAGATCAGAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 544
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
DB 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 604
QY 361 CGCTGTTTCTATGGCTTCAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
DB 605 CGCTGTTTCTATGGCTTCAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 480
DB 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 724
QY 481 ACCATGCCCTTANGTTAAGAAAAAGCAGATTTGGCCCTTGCATGAGATAGAGATAGAAAA 540
DB 725 ACCATGCCCTTANGTTAAGAAAAAGCAGATTTGGCCCTTGCATGAGATAGAGATAGAAAA 784
QY 541 TCTACTTTTGGGAAAGAGTGGGCTTTGCTGCTGTAGAGAGGTTTCTTCTCAGGA 600
DB 785 TCTACTTTTGGGAAAGAGTGGGCTTTGCTGCTGTAGAGAGGTTTCTTCTCAGGA 844
QY 601 TCTTTTGTCTATATCTGGCTTAAAGAAAGAGAGGTTCTATGCCAGGACTCACATTTTTC 660
DB 845 TCTTTTGTCTATATCTGGCTTAAAGAAAGAGAGGTTCTATGCCAGGACTCACATTTTTC 904
QY 661 AATGAACATCATCAGCAGATGAAGGACTTTCATGTGACTTTTGTGCTGATGTTCCAA 720
```

```
DB 905 AATGAACATCATCAGCAGATGAAGGACTTTCATGTGACTTTGCTTGCTGATGTTCCAA 964
QY 721 TACTTAGTAATAAGCCTTCAGAAAGAGGTCAGAGGATCATTTGTTGATGCTGTCAAA 780
DB 965 TACTTAGTAATAAGCCTTCAGAAAGAGGTCAGAGGATCATTTGTTGATGCTGTCAAA 1024
QY 781 ATTGAGCAGAGTTTAAACAGAGCCTTCCAGTTGGCCTCATTTGGAATCAATTTGCAAT 840
DB 1025 ATTGAGCAGAGTTTAAACAGAGCCTTCCAGTTGGCCTCATTTGGAATCAATTTGCAAT 1084
QY 841 TTGATGAAAACAGTACATTTAGTTTGTAGCTGACAGATTAATCTTGTGGAACCTTGGATTTCA 900
DB 1085 TTGATGAAAACAGTACATTTAGTTTGTAGCTGACAGATTAATCTTGTGGAACCTTGGATTTCA 1144
QY 901 AAGTCTTTTTCAGCAGAAAAATCCTTTTGAATTTATGAAAAACATTTCTTTAGAGGAAAA 960
DB 1145 AAGTCTTTTTCAGCAGAAAAATCCTTTTGAATTTATGAAAAACATTTCTTTAGAGGAAAA 1204
QY 961 ACAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTATATGCGCAGAAACC 1020
DB 1205 ACAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTATATGCGCAGAAACC 1264
QY 1021 ACAGATAACGCTTTCACCTTTGGATGCAGATTTT 1053
DB 1265 ACAGATAACGCTTTCACCTTTGGATGCAGATTTT 1297

RESULT 10
AR454877
LOCUS AR454877 1053 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 12 from patent US 6682917.
ACCESSION AR454877
VERSION AR454877.1 GI:42688832
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1053)
AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE Protein having a ribonucleotide Reductase activity and a DNA
theresof
JOURNAL Patent: US 6682917-A 12 27-JAN-2004;
FEATURES
source
1..1053
/mol_type="genomic DNA"

ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 1053;
Best Local Similarity 99.9%; Pred. No. 3e-255;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGACCCGGAAGCGGAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
DB 1 ATGGCGACCCGGAAGCGGAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 120
DB 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 120
QY 121 CGGTTTGTCATCTTTCATATCCAGTACCTGATATTTGGAAAAATGATATAACAGGCACAG 180
DB 121 CGGTTTGTCATCTTTCATATCCAGTACCTGATATTTGGAAAAATGATATAACAGGCACAG 180
QY 181 GCTTCTCTTGGACAGCAGAGAGTGCAGTATCAAGAGTCTCCCTCACCTGGAAACAG 240
DB 181 GCTTCTCTTGGACAGCAGAGAGTGCAGTATCAAGAGTCTCCCTCACCTGGAAACAG 240
QY 241 CTTAAAGCAGATGAGAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 300
DB 241 CTTAAAGCAGATGAGAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 300
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
```


Db 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGAGGTTCCAGAGGCT 360
Qy 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATGCCAAGAAAAGGAATTTTATTAATGCAATGAA 480
Db 421 CTGATAGACACTTACATCAGAGATGCCAAGAAAAGGAATTTTATTAATGCAATGAA 480
Qy 481 ACCATGCCCTATCTTAAGAAAAGAGAGATTTGGGCGCTTGGATAGGATAGCAGATAGAAA 540
Db 481 ACCATGCCCTATCTTAAGAAAAGAGAGATTTGGGCGCTTGGATAGGATAGCAGATAGAAA 540
Qy 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTTGGCTGTAGAAAGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTTGGCTGTAGAAAGAGTTTCTTCTCAGGA 600
Qy 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTTCTTATGCCAGGACTCATTTTTTC 660
Db 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTTCTTATGCCAGGACTCATTTTTTC 660
Qy 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Qy 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTTGAGGAGATCAATGTTGATGCTGTCAAA 780
Db 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTTGAGGAGATCAATGTTGATGCTGTCAAA 780
Qy 781 ATTGAGCAGGAGTTTAAAGAGGCTTGGCAGTTGGCTCAATGGAATGAAATGCAAT 840
Db 781 ATTGAGCAGGAGTTTAAAGAGGCTTGGCAGTTGGCTCAATGGAATGAAATGCAAT 840
Qy 841 TTGATGAAACAGTACATGAGTTTGTAGCTGACAGATTACTGTGGAACCTTGAATCTCA 900
Db 841 TTGATGAAACAGTACATGAGTTTGTAGCTGACAGATTACTGTGGAACCTTGAATCTCA 900
Qy 901 AAGTTTTCAGCAGAGAAATCTTTTGAATTTATGGAACAACTTTCTTTAGAGGAAA 960
Db 901 AAGTTTTCAGCAGAGAAATCTTTTGAATTTATGGAACAACTTTCTTTAGAGGAAA 960
Qy 961 ACAAATTTCTTTCAGAAACAGTTTCAGAGTATCAGGTTTTCAGTTATGCGAGAAACC 1020
Db 961 ACAAATTTCTTTCAGAAACAGTTTCAGAGTATCAGGTTTTCAGTTATGCGAGAAACC 1020
Qy 1021 ACAGATAACCTTCTTACCTTTGGATGAGATTTT 1053
Db 1021 ACAGATAACCTTCTTACCTTTGGATGAGATTTT 1053

RESULT 11
BD064773
LOCUS BD064773 1053 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel protein and DNA thereof.
ACCESSION BD064773
VERSION BD064773.1 GI:22610376
KEYWORDS JP 2001269184-A/10.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1053)
REFERENCE Nakamura, Y., Arakawa, H. and Tanaka, H.
AUTHORS Novel protein and DNA thereof
TITLE Patent: JP 2001269184-A 10 02-OCT-2001.
JOURNAL YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD
COMMENT OS Homo sapiens (human)
PN JP 2001269184-A/10
PD 02-OCT-2001
PE 27-JUN-2000 JP 2000192401
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC

C12N15/09, A61K31/711, A61K38/00, A61K38/43, A61K45/00, A61K48/00, PC
A61P35/00,
PC A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC
, C12N5/10, C12N9/02,
PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53// (C12N15/09,
C12R1:91),
PC (C12N15/09, C12R1:19), (C12N9/02, C12R1:91), (C12P21/02, C12R1:91),
PC (C12P21/08, C12R1:91), (C12N15/00, A61K37/02, A61K37/48, C12N5/00,
PC (C12N15/00, C12R1:91), (C12N15/00, C12R1:19)
CC Novel protein and DNA thereof
FH Key Location/Qualifiers
FT source 1..1053
FT /organism='Homo sapiens (human)'.
FEATURES
source
1..1053
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 1053;
Best Local Similarity 99.9%; Pred. No. 3e-255;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGGGCGACCCGGAAGCGCGGAGCGGCTGGATCAGATGAGATGAGATCATCTTCA 60
Db 1 ATGGGCGACCCGGAAGCGCGGAGCGGCTGGATCAGATGAGATGAGATCATCTTCA 60
Qy 61 GACACCAACCAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACCAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Qy 121 CGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAAAATGTATAAACAGGCACAG 180
Db 121 CGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAAAATGTATAAACAGGCACAG 180
Qy 181 GCTTCTCTTGGACAGAGAGGTCGACTTATCAAGAGATCTCCCTCAGTCGAGCAAG 240
Db 181 GCTTCTCTTGGACAGAGAGGTTGACTTATCAAGAGATCTCCCTCAGTCGAGCAAG 240
Qy 241 CTTAAGCAGATGAGAGTACTTCTATCTCAGATCTTAGCGCTTTTTCAGGCGAGTAT 300
Db 241 CTTAAGCAGATGAGAGTACTTCTATCTCAGATCTTAGCGCTTTTTCAGGCGAGTAT 300
Qy 301 GGAATTCCTAATGAAATTTGGTGGAGGCTTTAGTCAGAGAGTGCAGGCTCCAGAGCT 360
Db 301 GGAATTCCTAATGAAATTTGGTGGAGGCTTTAGTCAGAGAGTGCAGGCTCCAGAGCT 360
Qy 361 CGCTGTTTCTATGGCTTTCAAAATTTCTATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAAAATTTCTATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATGCCAAGAAAAGGAATTTTATTAATGCAATGAA 480
Db 421 CTGATAGACACTTACATCAGAGATGCCAAGAAAAGGAATTTTATTAATGCAATGAA 480
Qy 481 ACCATGCCCTATCTTAAGAAAAGAGAGATTTGGGCGCTTGGATAGGATAGCAGATAGAAA 540
Db 481 ACCATGCCCTATCTTAAGAAAAGAGAGATTTGGGCGCTTGGATAGGATAGCAGATAGAAA 540
Qy 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTTGGCTGTAGAAAGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTTGGCTGTAGAAAGAGTTTCTTCTCAGGA 600
Qy 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGGTTCTTATGCCAGGACTCATTTTTTC 660
Db 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGGTTCTTATGCCAGGACTCATTTTTTC 660
Qy 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Qy 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTTGAGGAGATCAATGTTGATGCTGTCAAA 780

```
Db 721 TACTTAGTAATAAGCCCTTCAGAAAGAGGTCAGGAGATCAATGTTGATGCTGTCAAA 780
Qy 781 ATTGACGAGAGTATTTTAAACAGAGGCTTGCAGTTGGCCTCAATGGAATGAATTTGCATT 840
Db 781 ATTGACGAGAGTATTTTAAACAGAGGCTTGCAGTTGGCCTCAATGGAATGAATTTGCATT 840
Qy 841 TTGATGAACACGATACATGATGAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTCTCA 900
Db 841 TTGATGAACACGATACATGATGAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTCTCA 900
Qy 901 AAGGTTTTCAGGACAGAAATCCTTTTGAATTTATGGAACATTTCTTTAGAGGAAA 960
Db 901 AAGGTTTTCAGGACAGAAATCCTTTTGAATTTATGGAACATTTCTTTAGAGGAAA 960
Qy 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
Qy 1021 ACAGATAACGTCCTTACCTTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGTCCTTACCTTTGGATGAGATTTT 1053

RESULT 12
BD093086
LOCUS
DEFINITION Novel protein and its DNA.
ACCESSION BD093086
VERSION BD093086.1 GI:22638674
KEYWORDS WO 0100799-A/10.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1053)
Nakamura,Y., Arakawa,H. and Tanaka.H.
Novel protein and its DNA
Patent: WO 0100799-A 10 04-JAN-2001;
TAKEDA CHEMICAL INDUSTRIES LTD,YUSUKE NAKAMURA,HIROFUMI ARAKAWA,
HIROSHI TANAKA
OS Homo sapiens (human)
PN WO 0100799-A/10
PD 04-JAN-2001
PF 27-JUN-2000 WO 2000JP004189
PR 28-JUN-1999 JP 99P 181131,06-JUL-1999 JP 99P 192391 PR
21-JAN-2000 JP 00P 017770
PI YUSUKE NAKAMURA,HIROFUMI ARAKAWA,HIROSHI TANAKA PC
C12N9/04,C12N15/53,C12N1/15,C12N1/19,C12N1/21,C12N5/10 PC
C12P21/02,A61K38/44,
PC A61K48/00,C07K16/40,G01N33/53,C12Q1/26
CC
FH Key Location/Qualifiers.
FEATURES
source 1..1053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 1053;
Best Local Similarity 99.9%; Pred. No. 3e-255;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGGGGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATGAGATGATCATCTTCA 60
Db 1 ATGGGGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATGAGATGATCATCTTCA 60
Qy 61 GACACCAACCAAGTGAATAAAGTCAATGAGAGCCACTCTCTAGAAAGCTTCTCGC 120
Db 61 GACACCAACCAAGTGAATAAAGTCAATGAGAGCCACTCTCTAGAAAGCTTCTCGC 120
Qy 121 CGGTTTGTGTCATCTTCCAAATCCAGTACCCTGATATTTGGAAATGTATAAACAGGCACAG 180
```

```
Db 121 CGGTTTGTGTCATCTTCCAAATCCAGTACCCTGATATTTGGAAATGTATAAACAGGCACAG 180
Qy 181 GCTTCTCTTCTGGACAGCAGAGAGGTCGACTTATCAAAGGATCTCCCTCACTGGAAACAAG 240
Db 181 GCTTCTCTTCTGGACAGCAGAGAGGTCGACTTATCAAAGGATCTCCCTCACTGGAAACAAG 240
Qy 241 CTTAAAGCAGATGAGAAAGTACTTCTCATCTCTCAATCTTACCTTTTGGAGCCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAAAGTACTTCTCATCTCTCAATCTTACCTTTTGGAGCCAGTGAT 300
Qy 301 GGAATTCGTAATGAAATTTTGGAGCGCTTTAGTCAGGAGGTCAGGTTCCAGAGGCT 360
Db 301 GGAATTCGTAATGAAATTTTGGAGCGCTTTAGTCAGGAGGTCAGGTTCCAGAGGCT 360
Qy 361 CGCTGTTTCTATGCTTTCAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGCTTTCAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACA CTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATGAA 480
Db 421 CTGATAGACA CTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATGAA 480
Qy 481 ACCATGCCCTATCTTAAGAAAGAGCAGATTTGGCCCTTGCATGATGATAGAGAA 540
Db 481 ACCATGCCCTATCTTAAGAAAGAGCAGATTTGGCCCTTGCATGATGATAGAGAA 540
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTCTCTCTGTAGAAAGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTCTCTCTGTAGAAAGAGTTTCTTCTCAGGA 600
Qy 601 TCTTTTGTCTGCTATATTCTGGCTAAAGAGAGGTCTTATGCGAGACTCACTTTTTC 660
Db 601 TCTTTTGTCTGCTATATTCTGGCTAAAGAGAGGTCTTATGCGAGACTCACTTTTTC 660
Qy 661 AATGAATCTATCAGCAGAGATGAAGACTTCACTGTGACTTGTGCTGCTGATGTTCCAA 720
Db 661 AATGAATCTATCAGCAGAGATGAAGACTTCACTGTGACTTGTGCTGCTGATGTTCCAA 720
Qy 721 TACTTAGTAATAAGCCTTCAGAAAGAGGTCAGGAGATCAATGTTGATGCTGTCAAA 780
Db 721 TACTTAGTAATAAGCCTTCAGAAAGAGGTCAGGAGATCAATGTTGATGCTGTCAAA 780
Qy 781 ATTGAGCAGGAGTATTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATTCATT 840
Db 781 ATTGAGCAGGAGTATTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATTCATT 840
Qy 841 TTGATGAACACGATACATGATGAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTCTCA 900
Db 841 TTGATGAACACGATACATGATGAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTCTCA 900
Qy 901 AAGGTTTTCAGGACAGAAATCCTTTTGAATTTATGGAACATTTCTTTAGAGGAAA 960
Db 901 AAGGTTTTCAGGACAGAAATCCTTTTGAATTTATGGAACATTTCTTTAGAGGAAA 960
Qy 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
Qy 1021 ACAGATAACGTCCTTACCTTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGTCCTTACCTTTGGATGAGATTTT 1053

RESULT 13
AR454870
LOCUS
DEFINITION Sequence 4 from patent US 6682917.
ACCESSION AR454870
VERSION AR454870.1 GI:42688825
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
```

Unclassified.
1 (bases 1 to 1081)
Nakamura, Y., Arakawa, H. and Tanaka, H.
Protein having a ribonucleotide Reductase activity and a DNA
thereof
Patent: US 682917-A 4 27-JAN-2004;
JOURNAL Location/Qualifiers
FEATURES
source 1..1081
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 1081;
Best Local Similarity 99.9%; Pred. No. 3e-255;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGGCGACCGGAAAGCGCGGAGCGGCTGGATCAGATGAGAGATCATCTTCA 60
DB 20 ATGGGCGACCGGAAAGCGGAGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 79
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 80 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 139
QY 121 CGGTTTGTCATCTTCCAAATCCAGTACCTGTATTTGGAAATGTATAACAGGCACAG 180
DB 140 CGGTTTGTCATCTTCCAAATCCAGTACCTGTATTTGGAAATGTATAACAGGCACAG 199
QY 181 GCTTCCTCTGGACAGCAGAGAGTGCATTTATCAAGAGATCTCCCTCACTGGAACAAG 240
DB 200 GCTTCCTCTGGACAGCAGAGAGTGCATTTATCAAGAGATCTCCCTCACTGGAACAAG 259
QY 241 CTTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCTTTTGGCAGCAGTGAT 300
DB 260 CTTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCTTTTGGCAGCAGTGAT 319
QY 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAAGTTCCAGAGGCT 360
DB 320 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAAGTTCCAGAGGCT 379
QY 361 CGCTGTTTCTATGGCTTCAAAATTTCTCATCGAAGTGTTCACTCAGAGATGTACAGTTTG 420
DB 380 CGCTGTTTCTATGGCTTCAAAATTTCTCATCGAAGTGTTCACTCAGAGATGTACAGTTTG 439
QY 421 CTGATAGACATTTACATCAGATGCCAAGAAAGGAAATTTTATTAATGCAATTTGAA 480
DB 440 CTGATAGACATTTACATCAGATGCCAAGAAAGGAAATTTTATTAATGCAATTTGAA 499
QY 481 ACCATGCCCTATGTTAAGAAAGAGATTTGGGCGCTTGGATGAGATAGAGAA 540
DB 500 ACCATGCCCTATGTTAAGAAAGAGATTTGGGCGCTTGGATGAGATAGAGAA 559
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCGCTTGTCTGTAGAGAGTGTCTTCTCAGGA 600
DB 560 TCTACTTTTGGGAAAGAGTGGTGGCGCTTGTCTGTAGAGAGTGTCTTCTCAGGA 619
QY 601 TCTTTTGTCTATATCTGGCTTAAGAAAGAGTGTCTTATGCCAGACTCATTTTTC 660
DB 620 TCTTTTGTCTATATCTGGCTTAAGAAAGAGTGTCTTATGCCAGACTCATTTTTC 679
QY 661 AATGAATCATCAGCAGATGAAGACTTTCATCTGACATTTTGTCTGCTGATGTTCCAA 720
DB 680 AATGAATCATCAGCAGATGAAGACTTTCATCTGACATTTTGTCTGCTGATGTTCCAA 739
QY 721 TACTTAGTAAATAGCCCTTCAGAAAGAGGTTGAGGAGATCATTTGTGATGCTGTCAAA 780
DB 740 TACTTAGTAAATAGCCCTTCAGAAAGAGGTTGAGGAGATCATTTGTGATGCTGTCAAA 799
QY 781 ATTGAGCAGAGTGTAAAGAGCCCTTGCAGTGTGGCCCTCATTTGGAATGAAATTCAT 840
DB 800 ATTGAGCAGAGTGTAAAGAGCCCTTGCAGTGTGGCCCTCATTTGGAATGAAATTCAT 859
QY 841 TTGATGAAACAGTACATTTGAGTTGTAGCTGACAGATTTACTTGTGGAACTTGGATTCTCA 900

Db 860 TTGATGAAACAGTACATTTGAGTTGTAGCTGACAGATTTACTTGTGGAACTTGGATTCTCA 919
QY 901 AAGGTTTTTTCAGGCAGAAAAATCCTTTTGTATTTATGGAAAAATTTCTTTTAGAAGAAAA 960
DB 920 AAGGTTTTTTCAGGCAGAAAAATCCTTTTGTATTTATGGAAAAATTTCTTTTAGAAGAAAA 979
QY 961 ACAAATTTCTTTTGAGAAACGAGTTTCAGAGATATCAGCGTTTTCAGTTATGGCAGAAACC 1020
DB 980 ACAAATTTCTTTTGAGAAACGAGTTTCAGAGATATCAGCGTTTTCAGTTATGGCAGAAACC 1039
QY 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053
DB 1040 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1072

RESULT 14
BD064766
LOCUS Novel protein and DNA thereof.
DEFINITION
ACCESSION BD064766
VERSION BD064766.1 GI:22610369
KEYWORDS JP 2001269184-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1081)
AUTHORS Nakamura, Y., Arakawa, H. and Tanaka, H.
TITLE Novel protein and DNA thereof
JOURNAL Patent: JP 2001269184-A 3 02-OCT-2001;
YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD
COMMENT OS Homo sapiens (human)
PN JP 2001269184-A/3
PD 02-OCT-2001
PF 27-JUN-2000 JP 2000192401
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
A61P35/00, A61K31/711, A61K38/00, A61K38/43, A61K45/00, A61K48/00, PC
A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC
, C12N5/10, C12N9/02,
PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53// (C12N15/09,
PC C12R1:91),
PC (C12N15/09, C12R1:19), (C12N9/02, C12R1:91), (C12P21/02, C12R1:91),
PC (C12P21/08, C12R1:91), C12N15/00, A61K37/02, A61K37/48, C12N5/00,
PC (C12N15/00, C12R1:91), (C12N15/00, C12R1:19)
CC Novel protein and DNA thereof
FH Key Location/Qualifiers
FT source 1..1081
FT location/Qualifiers
1..1081
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source 1..1081
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 1081;
Best Local Similarity 99.9%; Pred. No. 3e-255;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGGCGACCGGAAAGCGCGGAGCGGCTGGATCAGATGAGAGATCATCTTCA 60
DB 20 ATGGGCGACCGGAAAGCGGAGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 79
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 80 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 139
QY 121 CGGTTTGTCATCTTCCAAATCCAGTACCTGTATTTGGAAATGTATAACAGGCACAG 180
DB 140 CGGTTTGTCATCTTCCAAATCCAGTACCTGTATTTGGAAATGTATAACAGGCACAG 199
QY 181 GCTTCCTCTGGACAGCAGAGAGTGCATTTATCAAGAGATCTCCCTCACTGGAACAAG 240
DB 200 GCTTCCTCTGGACAGCAGAGAGTGCATTTATCAAGAGATCTCCCTCACTGGAACAAG 259
QY 241 CTTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCTTTTGGCAGCAGTGAT 300
DB 260 CTTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCTTTTGGCAGCAGTGAT 319
QY 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAAGTTCCAGAGGCT 360
DB 320 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAAGTTCCAGAGGCT 379
QY 361 CGCTGTTTCTATGGCTTCAAAATTTCTCATCGAAGTGTTCACTCAGAGATGTACAGTTTG 420
DB 380 CGCTGTTTCTATGGCTTCAAAATTTCTCATCGAAGTGTTCACTCAGAGATGTACAGTTTG 439
QY 421 CTGATAGACATTTACATCAGATGCCAAGAAAGGAAATTTTATTAATGCAATTTGAA 480
DB 440 CTGATAGACATTTACATCAGATGCCAAGAAAGGAAATTTTATTAATGCAATTTGAA 499
QY 481 ACCATGCCCTATGTTAAGAAAGAGATTTGGGCGCTTGGATGAGATAGAGAA 540
DB 500 ACCATGCCCTATGTTAAGAAAGAGATTTGGGCGCTTGGATGAGATAGAGAA 559
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCGCTTGTCTGTAGAGAGTGTCTTCTCAGGA 600
DB 560 TCTACTTTTGGGAAAGAGTGGTGGCGCTTGTCTGTAGAGAGTGTCTTCTCAGGA 619
QY 601 TCTTTTGTCTATATCTGGCTTAAGAAAGAGTGTCTTATGCCAGACTCATTTTTC 660
DB 620 TCTTTTGTCTATATCTGGCTTAAGAAAGAGTGTCTTATGCCAGACTCATTTTTC 679
QY 661 AATGAATCATCAGCAGATGAAGACTTTCATCTGACATTTTGTCTGCTGATGTTCCAA 720
DB 680 AATGAATCATCAGCAGATGAAGACTTTCATCTGACATTTTGTCTGCTGATGTTCCAA 739
QY 721 TACTTAGTAAATAGCCCTTCAGAAAGAGGTTGAGGAGATCATTTGTGATGCTGTCAAA 780
DB 740 TACTTAGTAAATAGCCCTTCAGAAAGAGGTTGAGGAGATCATTTGTGATGCTGTCAAA 799
QY 781 ATTGAGCAGAGTGTAAAGAGCCCTTGCAGTGTGGCCCTCATTTGGAATGAAATTCAT 840
DB 800 ATTGAGCAGAGTGTAAAGAGCCCTTGCAGTGTGGCCCTCATTTGGAATGAAATTCAT 859
QY 841 TTGATGAAACAGTACATTTGAGTTGTAGCTGACAGATTTACTTGTGGAACTTGGATTCTCA 900

[illegible]

Search completed: June 16, 2005, 08:31:35
Job time : 4893 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2005, 06:56:56 ; Search time 220 Seconds
(without alignments)
7831.815 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1053

Sequence: 1 atggcgaccggaagcc.....tcaccttgatgagagatttt 1053

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PTCUS COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1053	100.0	1053	4	US-10-019-733-2
2	1053	100.0	4955	4	US-10-019-733-3
3	1051.4	99.8	1053	4	US-10-019-733-12
4	1051.4	99.8	1081	4	US-10-019-733-4
5	579	55.0	2500	4	US-09-962-665-9
6	579	55.0	2500	4	US-09-962-677-9
7	579	55.0	2500	4	US-09-962-677-9
8	578.4	54.9	2479	4	US-09-949-016-2025
9	578.4	54.9	2500	4	US-09-023-655-1370
10	578.4	54.9	2500	4	US-09-949-016-145
11	441	41.9	481	3	US-08-905-223-125
12	429.8	40.8	14176	1	US-08-307-499-1
13	429.8	40.8	14176	1	US-08-307-499-14
14	429.8	40.8	14176	3	US-09-299-268-1
15	429.8	40.8	14176	3	US-09-299-268-14
16	383.2	36.4	1236	4	US-09-248-796A-3914
17	382.8	36.4	825	1	US-08-307-499-40
18	382.8	36.4	825	3	US-09-299-268-40
19	337.8	32.1	1112	1	US-08-136-743B-1
20	318.2	30.2	419	3	US-08-905-223-129
21	316	30.0	1083	4	US-09-248-796A-3915
22	260.8	24.8	656	3	US-09-328-111-791
23	195.2	18.5	449	4	US-09-270-767-14670
24	169.2	16.1	366	3	US-08-307-499-13
25	169.2	16.1	366	3	US-09-299-268-13
26	141.4	13.4	373	4	US-09-401-064-337
27	139.8	13.3	373	4	US-09-401-064-319

C	28	115.2	10.9	366	4	US-09-401-064-307	Sequence 307, App
	29	112.4	10.7	297	4	US-09-313-294A-4435	Sequence 4435, Ap
	30	93.8	8.9	11820	4	US-09-949-016-11887	Sequence 11887, A
	31	93.8	8.9	11826	4	US-09-949-016-13767	Sequence 13767, A
	32	92	8.7	601	4	US-09-949-016-20615	Sequence 20615, A
	33	92	8.7	601	4	US-09-949-016-69028	Sequence 69028, A
	34	74.4	7.1	276	4	US-09-313-294A-3437	Sequence 3437, Ap
	35	70.2	6.7	304	4	US-09-313-294A-7143	Sequence 7143, Ap
	36	68	6.5	276	4	US-09-313-294A-4638	Sequence 4638, Ap
	37	61	5.8	1053	4	US-09-902-540-2666	Sequence 2666, Ap
	38	61	5.8	13706	4	US-09-902-540-1124	Sequence 1124, Ap
	39	60	5.7	601	4	US-09-949-016-69031	Sequence 69031, A
	40	59.8	5.7	1141	4	US-09-806-708B-22	Sequence 22, Appl
	41	54.4	5.2	288	4	US-09-313-294A-6108	Sequence 6108, Ap
	42	52.2	5.0	276	4	US-09-313-294A-3244	Sequence 3244, Ap
	43	50	4.7	1230025	4	US-09-198-452A-1	Sequence 1, Appli
	44	50	4.7	1230230	4	US-09-438-185A-1	Sequence 1, Appli
	45	45.4	4.3	124884	4	US-09-661-598A-76	Sequence 76, Appl

ALIGNMENTS

RESULT 1

US-10-019-733-2
; Sequence 2, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-2

Query Match	100.0%;	Score 1053;	DB 4;	Length 1053;
Best Local Similarity	100.0%;	Pred. No. 8.7e-311;		
Matches 1053;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGCGCACCCGGAAGCGCGGCGGCTGGATCAGGATGAGATCATCTTCA	60	
Db	1	ATGGCGCACCCGGAAGCGCGGCGGCTGGATCAGGATGAGATCATCTTCA	60	
QY	61	GACACCAAGAAAGTGAATAAAGTCAATGAAGCCACTCTTAAGAAAGATTTCTCG	120	
Db	61	GACACCAAGAAAGTGAATAAAGTCAATGAAGCCACTCTTAAGAAAGATTTCTCG	120	
QY	121	CGGTTTGTCATCTTTCCCAATCCAGTACCTGTATTTGAAAATGTATAAACAGGCACAG	180	
Db	121	CGGTTTGTCATCTTTCCCAATCCAGTACCTGTATTTGAAAATGTATAAACAGGCACAG	180	
QY	181	GCTTCCTTCTTGACACGACGAGAGTTCGACTTATCAAGAGATCTCCCTCACTGGAACAAG	240	
Db	181	GCTTCCTTCTTGACACGACGAGAGTTCGACTTATCAAGAGATCTCCCTCACTGGAACAAG	240	
QY	241	CTTAAAGCAGATGAGAAGTACTTCTCTCATCTTACCTTTTTCGAGCCAGTGAT	300	
Db	241	CTTAAAGCAGATGAGAAGTACTTCTCTCATCTTACCTTTTTCGAGCCAGTGAT	300	
QY	301	GGAATTGTAATGAAATTTGGTGGCGGCTTTAGTCAGGAGGTGCGAGGTTCAGAGGCT	360	

```

Db 301 GGAATTTGTAATCAAAATTTGTTGGAGCGCTTTAGTCAGAGGTGAGGTTTCAGAGGCT 360
Qy 361 CCCTGTTTCTATGGCTTTCAAAATTTCTATCGAGATGTTTCACTCAGAGATGTACAGTTTG 420
Db 361 CCCTGTTTCTATGGCTTTCAAAATTTCTATCGAGATGTTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACATTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGAA 480
Db 421 CTGATAGACATTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAGAGAGATGAGGCTTGGCTGCTGAGAGAGTTTCTCTCAGGA 540
Db 481 ACCATGCCCTATGTTAAGAAAGAGAGATGAGGCTTGGCTGCTGAGAGATGAGAGAA 540
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGAGAGAGTTTCTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGAGAGAGTTTCTCTCAGGA 600
Qy 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGTCTTATGCGAGACTCACTTTTTC 660
Db 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGTCTTATGCGAGACTCACTTTTTC 660
Qy 661 AATGAACCTCATCAGCAGAGATGAAGACTTCTACTGTGACTTTGCTGCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGACTTCTACTGTGACTTTGCTGCTGATGTTCCAA 720
Qy 721 TACTTAGTAAATAAGCTTTCAGAGAAAGGCTCAGGAGAGATCATTTGATGCTGTCAAA 780
Db 721 TACTTAGTAAATAAGCTTTCAGAGAAAGGCTCAGGAGAGATCATTTGATGCTGTCAAA 780
Qy 781 ATTGAGCAGGAGTTTTTAAAGAGAGCTTTGCCAGTGGCTCATTTGGAATGAAATTCATT 840
Db 781 ATTGAGCAGGAGTTTTTAAAGAGAGCTTTGCCAGTGGCTCATTTGGAATGAAATTCATT 840
Qy 841 TTGATGAACACGTACATTGAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTTCA 900
Db 841 TTGATGAACACGTACATTGAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTTCA 900
Qy 901 AAGGTTTTTCAGCAGAGAAATCTTTTGAATTTTATGGAAGAAATTTCTTTAGAGGAA 960
Db 901 AAGGTTTTTCAGCAGAGAAATCTTTTGAATTTTATGGAAGAAATTTCTTTAGAGGAA 960
Qy 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGATCAGCGTTTTCAGCTTATGCGAGAAAC 1020
Db 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGATCAGCGTTTTCAGCTTATGCGAGAAAC 1020
Qy 1021 ACAGATAACGTTCTTCACTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGTTCTTCACTTGGATGAGATTTT 1053

```

RESULT 2

```

US-10-019-733-3
; Sequence 3, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 3
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens

```

```

; FEATURE:
; OTHER INFORMATION:
US-10-019-733-3
Query Match 100.0%; Score 1053; DB 4; Length 4955;
Best Local Similarity 100.0%; Pred. No. 2.1e-310;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGGCGACCCGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 245 ATGGGCGACCCGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
Qy 61 GACACCAACCAAGTGAATAAAGTCAANTGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 305 GACACCAACCAAGTGAATAAAGTCAANTGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
Qy 121 CGGTTTGTCTATCTTTCCAAATCCAGTACCTGATATTTGGAATAATGTATAAAGAGCAG 180
Db 365 CGGTTTGTCTATCTTTCCAAATCCAGTACCTGATATTTGGAATAATGTATAAAGAGCAG 424
Qy 181 GCTTCTCTTCTGGACAGAGAGTGCAGTTATCAAGAGATCTCCCTCACTGGAAACAAG 240
Db 425 GCTTCTCTTCTGGACAGAGAGTGCAGTTATCAAGAGATCTCCCTCACTGGAAACAAG 484
Qy 241 CTTAAGCAGATGAGAAAGTACTTCTCATCTCAGATCTTAGCCTTTTTCAGGCGAGTGT 300
Db 485 CTTAAGCAGATGAGAAAGTACTTCTCATCTCAGATCTTAGCCTTTTTCAGGCGAGTGT 544
Qy 301 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGGTTCCAGAGGT 360
Db 545 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGGTTCCAGAGGT 604
Qy 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGATGTTTCACTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGATGTTTCACTCAGAGATGTACAGTTTG 664
Qy 421 CTGATAGACATTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGAA 480
Db 665 CTGATAGACATTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGAA 724
Qy 481 ACCATGCCCTATGTTAAGAAAGAGAGTGGGCTTTGCCAGTGGATGAGAGATGAGAA 540
Db 725 ACCATGCCCTATGTTAAGAAAGAGAGTGGGCTTTGCCAGTGGATGAGAGATGAGAA 784
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGAGTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGAGTTTCTTCTCAGGA 844
Qy 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGTCTTATGCGAGACTCACTTTTTC 660
Db 845 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGTCTTATGCGAGACTCACTTTTTC 904
Qy 661 AATGAACCTCATCAGCAGAGATGAGGACTTCACTGTGACTTTTTCGCTGATGTTCCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAGGACTTCACTGTGACTTTTTCGCTGATGTTCCAA 964
Qy 721 TACTTAGTAAATAAGCTTTCAGAGAAAGGCTCAGGAGAGATCATTTGATGCTGTCAAA 780
Db 965 TACTTAGTAAATAAGCTTTCAGAGAAAGGCTCAGGAGAGATCATTTGATGCTGTCAAA 1024
Qy 781 ATTGAGCAGGAGTTTTTAAAGAGAGCTTGCAGTTGCGCTCATTTGGAATGAAATTCATT 840
Db 1025 ATTGAGCAGGAGTTTTTAAAGAGAGCTTGCAGTTGCGCTCATTTGGAATGAAATTCATT 1084
Qy 841 TTGATGAACACGTACATTGAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTTCA 900
Db 1085 TTGATGAACACGTACATTGAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTTCA 1144
Qy 901 AAGGTTTTTCAGCAGAGAAATCTTTTGAATTTTATGGAAGAAATTTCTTTAGAGGAA 960
Db 1145 AAGGTTTTTCAGCAGAGAAATCTTTTGAATTTTATGGAAGAAATTTCTTTAGAGGAA 1204
Qy 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGATCAGCGCTTTTTCAGAGTATGCGAGAAAC 1020

```


Db 1205 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTGGAGTTATGGCAGAACCC 1264
QY 1021 ACAGATAAGCTCTTCCACCTTGGATGAGATTTT 1053
Db 1265 ACAGATAAGCTCTTCCACCTTGGATGAGATTTT 1297

RESULT 3

US-10-019-733-12
; Sequence 12, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WO0P
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-12

Query Match 99.8%; Score 1051.4; DB 4; Length 1053;
Best Local Similarity 99.9%; Pred. No. 2.7e-310;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGGAGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGGCGACCCGGAAGCGGAGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACCAAGAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAAGAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 121 CGGTTTGTCTATCTTTCCATCCAGTACCTCTGATATTGGAAATGTATAACAGGCACAG 180
Db 121 CGGTTTGTCTATCTTTCCATCCAGTACCTCTGATATTGGAAATGTATAACAGGCACAG 180
QY 181 GCTTCTCTCTGGACACAGAGAGTTCGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
Db 181 GCTTCTCTCTGGACACAGAGAGTTCGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCTCACTCTTAGCCCTTTTTCAGCCAGTAT 300
Db 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCTCACTCTTAGCCCTTTTTCAGCCAGTAT 300
QY 301 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Db 301 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTTTCAATTTCTCATCGAGAATGTTTCACATCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTTCAATTTCTCATCGAGAATGTTTCACATCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCCCTTGGATGAGATAGAGATAAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCCCTTGGATGAGATAGAGATAAAA 540

QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTCTGCTGTAGAGAGAGTTTCTTCTTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTCTGCTGTAGAGAGAGTTTCTTCTTCAGGA 600
QY 601 TCTTTTGGCTGTATATCTTGTCTAAAGAGAGAGGTCTTATGCCAGGAGCTCACTTTTCC 660
Db 601 TCTTTTGGCTGTATATCTTGTCTAAAGAGAGAGGTCTTATGCCAGGAGCTCACTTTTCC 660
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGTGCTTTGCTTGGCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGTGCTTTGCTTGGCTGATGTTCCAA 720
QY 721 TACTTAGTAAATAGCCCTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAAA 780
Db 721 TACTTAGTAAATAGCCCTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAAA 780
QY 781 ATTGAGCAGGAGTTTAAACAGAGAGCCCTTGGCCAGTTGGCCCTCAITGGGAATGAATTCATT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGAGCCCTTGGCCAGTTGGCCCTCAITGGGAATGAATTCATT 840
QY 841 TTGATGAACACAGTACATTTGAGTTTGTAGCTGACAGATTACTTTGTGGAACCTTGGATTCTCA 900
Db 841 TTGATGAACACAGTACATTTGAGTTTGTAGCTGACAGATTACTTTGTGGAACCTTGGATTCTCA 900
QY 901 AAGTTTTTTCAGGAGAGAAATCCTTTTGAATTTATGGAAGAAACATTTCTTTAGAGAGAAA 960
Db 901 AAGTTTTTTCAGGAGAGAAATCCTTTTGAATTTATGGAAGAAACATTTCTTTAGAGAGAAA 960
QY 961 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCGAGTTATGCGAGAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCGAGTTATGCGAGAAACC 1020
QY 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053

RESULT 4

US-10-019-733-4
; Sequence 4, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WO0P
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-4

Query Match 99.8%; Score 1051.4; DB 4; Length 1081;
Best Local Similarity 99.9%; Pred. No. 2.7e-310;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGGCGACCCGGAAGCGGAGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 20 ATGGGCGACCCGGAAGCGGAGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 79
QY 61 GACACCAAGAGTGAATTAAGTCAATTAAGAGAGCCACTCTTAAAGAAAGAGTTCTCGC 120
Db 80 GACACCAAGAGTGAATTAAGTCAATTAAGAGAGCCACTCTTAAAGAAAGAGTTCTCGC 139

```

QY 121 CGGTTTGTCACTTTTCCAAATCCAGTACCTCTGATATTTGGAAAATGTATAAACAGGCACAG 180
Db 140 CGGTTTGTCACTTTTCCAAATCCAGTACCTCTGATATTTGGAAAATGTATAAACAGGCACAG 199
QY 181 GCTTCCTTTGGACAGCAGAGAGGTGCTGATATCAAGGATCTCCTCACTGGAACAAG 240
Db 200 GCTTCCTTTGGACAGCAGAGAGGTGCTGATATCAAGGATCTCCTCACTGGAACAAG 259
QY 241 CTTAAAGCAGATCAGAAAGTACTTCACTCTCAATCTTACGCTTTTGGACGCAAGTAT 300
Db 260 CTTAAAGCAGATCAGAAAGTACTTCACTCTCAATCTTACGCTTTTGGACGCAAGTAT 319
QY 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTTCCAGAGGCT 360
Db 320 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTTCCAGAGGCT 379
QY 361 CGCTGTTTCTATGGCTTTCAAATCTCATCGAAGTGTTCACATCAGAGATGTACAGTTTG 420
Db 380 CGCTGTTTCTATGGCTTTCAAATCTCATCGAAGTGTTCACATCAGAGATGTACAGTTTG 439
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTAATGCAATGAA 480
Db 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTAATGCAATGAA 499
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCCCTTGGATGGATAGAGATAGAAAA 540
Db 500 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCCCTTGGATGGATAGAGATAGAAAA 559
QY 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTGGCTGTGATAGAGAGTTCCTCTCAGGA 600
Db 560 TCTACTTTTGGGAAAAGAGTGGTGGCTTGGCTGTGATAGAGAGTTCCTCTCAGGA 619
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGTCTTATGCGAGAGTCTTATGCGAGTCTTTC 660
Db 620 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGTCTTATGCGAGAGTCTTATGCGAGTCTTTC 679
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGCCTTGCCTGATGTCCAA 720
Db 680 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGCCTTGCCTGATGTCCAA 739
QY 721 TACTTAGTAAATGAAGCTTCAGAAAGAGGTGAGGAGATCAATTTGTGATGCTGTCAAA 780
Db 740 TACTTAGTAAATGAAGCTTCAGAAAGAGGTGAGGAGATCAATTTGTGATGCTGTCAAA 799
QY 781 ATTGACGAGAGTATTTTAAAGAGAGTCTGAGTGGCTTGGCTCATTTGGAATGAATTCATT 840
Db 800 ATTGACGAGAGTATTTTAAAGAGAGTCTGAGTGGCTTGGCTCATTTGGAATGAATTCATT 859
QY 841 TTGATGAAAACAGTACATTTGAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTTCA 900
Db 860 TTGATGAAAACAGTACATTTGAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTTCA 919
QY 901 AAGGTTTTTTCAGCAGAAAATCCTTTTGAATTTTATGAAAACATTTCTTTAGAGGAAAA 960
Db 920 AAGGTTTTTTCAGCAGAAAATCCTTTTGAATTTTATGAAAACATTTCTTTAGAGGAAAA 979
QY 961 ACAAAATTTCTTGAAGAACAGTTCAGATATCAGGTTTTCAGGTTTTCAGGTTTTCAGGTTTTC 1020
Db 980 ACAAAATTTCTTGAAGAACAGTTCAGATATCAGGTTTTCAGGTTTTCAGGTTTTCAGGTTTTC 1039
QY 1021 ACAGATAACGCTTTCACCTTTGGATGCAATTTT 1053
Db 1040 ACAGATAACGCTTTCACCTTTGGATGCAATTTT 1072

```

RESULT 5

US-09-962-665-9
; Sequence 9, Application US/09962665
; Patent No. 6537759
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: POLYLPOLYGLUTAMATE SYNTHETASE GENE SEQUENCE

```

; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING THE
; TITLE OF INVENTION: TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015004
; CURRENT APPLICATION NUMBER: US/09/962.665
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128..1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc_feature
; LOCATION: 1636..1738, 2259
; OTHER INFORMATION: n = c or t
; OTHER INFORMATION: n = c or t
US-09-962-665-9

```

Query Match 55.0%; Score 579; DB 4; Length 2500;

Best Local Similarity 75.0%; Pred. No. 6.4e-166; Mismatches 241; Indels 0; Gaps 0;

Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

```

QY 89 ATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGCCGGTTTGTCTCTTTCCTTCCAAATCCAGTACC 148
Db 397 AGGATGAGCCGCTGCTCAGAGAAAACCCCGCCGCTTTGTCTCTTCCCATCGAGTACC 456
QY 149 CTGATATTTGGAATAATGTATAAACAGGCAACAGGCTTCCCTCTTGACAGCAGAGAGGTCG 208
Db 457 ATGATATCTGGCAGATGTATAAGAAAGGCAGAGGCTTCCCTTTGGACCCGCGAGGAGTTG 516
QY 209 ACTTATCAAGAGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAAGTACTTCACTCT 268
Db 517 ACCTCTCNAAGGACATTCAGCAGTGGGAATCCCTTGAACCCGAGAGAGATATTTTATAT 576
QY 269 CTCACTATCTAGCCCTTTTTCAGCCAGTGAATGTAATGTAATGTAATGTAATGTAATGTAATG 328
Db 577 CCCATGTTCTGGCTTTCTTTTCAGCAGAGCGATGCAATGTAATGTAATGTAATGTAATGTAATG 636
QY 329 GCTTTAGTCAGAGGTCAGAGGTCAGAGGTCAGAGGTCAGAGGTCAGAGGTCAGAGGTCAGAGG 388
Db 637 GATTAGCAAGAGTTTCAAGATTACAGAAAGCCGCTGTTCTTATGCTTCCAAATTTGCCA 696
QY 389 TCGAGATGTTCACTCAGAGATGTACAGTTTGTCTGTAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAAAACATACATTCCTGAAATGTATAGTCTTCTTATGACACTTTACATAAAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTTAAATCAATGAAACATGCCCCTATGTTAAGAAAAAGCAG 508
Db 757 AAGAAAGGGAATTTCTCTTCAATGCCATTTGAAACGATGCCCTTGTCTCAAGAGAGGAGCAG 816
QY 509 ATTGGGCTTTCGAGTGGATAGCAGATAGAAATCTACTTTTGGGGAAGAGAGTGGTGCT 568

```

```
Db 817 ACTGGGCTTGCCTGATGGGACAAAGAGGCTACTATGTTGAAGCTGTTGACGCT 876
Qy 569 TTGCTGCTGTAGAGGAGTTTCTTCTCAGAGTCTTTTGTGCTATATTCCTGGCTAAGA 628
Db 877 TTGCTGAGTGAAGGCAATTTCTTTTCGCGTCTTTTGGCTCGATATTCCTGGCTCAAGA 936
Qy 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCATGAACTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGAGTGTGCTGGCTCCTCACAATTTCTTAATGAAGTATATAGCAGAGATGAGGTT 996
Qy 689 TTCACTGTGACTTTGCTTGTGCTGATGTTCCAACTATTAGTAATAAGCCTTTCAGAGAAA 748
Db 997 TACACTGTGATTTTGTGCTGCTGATGTTCAACACCTGGTACACAAACCATCGAGGAGA 1056
Qy 749 GGCTCAGGAGATCAITTTGTGCTGTCAAAATTTGAGCAGGAGTTTTTAAACGAAGCCT 808
Db 1057 GAGTAAGAGAAATAATATCAATGCTGTTCGGATAGAACAGGAGTTCCTCACTGAGGCT 1116
Qy 809 TGCAGTTGGCCTCATTTGGAATGAATTTGCAATTTTGAATGAAGCACTGATGTTGTAG 868
Db 1117 TGCCTGTGAAGCTCAITTTGGGATGAATTTGCACTCTAATGAAGCAATACATTTGAGTTGTGG 1176
Qy 869 CTGACAGATTACTTTGGAACTTGGATTCTCAAGGTTTTTCAGGAGAAATCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGAACTGGGTTTTTAGCAAGGTTTTTCAGAGTAGAGAACCCATTTG 1236
Qy 929 ATTTTATGAAAAATTTCTTTAGAGGAAAAACAAATTTCTTTTACCTTGGATGCTG 1356
Db 1049 ATTT 1052
Db 1357 ACTT 1360
```

RESULT 6

```
US-09-963-333-9
; Sequence 9, Application US/09963333
; Patent No. 6664062
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE SYNTHASE GENE SEQUENCE VARIANCES
; TITLE OF INVENTION: HAVING UTILITY IN DETERMINING THE TREATMENT
; FILE REFERENCE: 11926-015002
; CURRENT APPLICATION NUMBER: US/09/963,333
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128, 1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
```

```
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc_feature
; LOCATION: 1636, 1738, 2259
; OTHER INFORMATION: n = c or t
US-09-963-333-9
```

Query Match 55.0%; Score 579; DB 4; Length 2500;

Best Local Similarity 75.0%; Pred. No. 6.4e-166;

Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

```
Qy 89 ATGAGAGCCACTCCTAGAAAGAGTTCTCGCCGCTTGTCTCATCTTTCCAAATCCAGTACC 148
Db 397 AGGATGAGCCCTGCTGAGAGAAAACCCCGCGCTTGTCTCATCTTTCCCATCGAGTACC 456
Qy 149 CTGATATTTGAAAAATGTATAACAGGCACAGGCTTCTCTTGGACAGCAGAGAGGCTCG 208
Db 457 ATGATATCTGCAGATGTATAAGAGGCGAGGCTTCTTTTGGACCGCGGAGGAGTTG 516
Qy 209 ACTTATCAAGAGATCTCCCTCACTGGAACAAAGCTTTAAAGCAGATGAGAGTACTTCATCT 268
Db 517 ACCTCTNAAGGACATTGAGCACTGGGAATCCCTGAAACCCCGAGGAGAGATATTTTATAT 576
Qy 269 CTCACTATTAGCCCTTTTTCAGCCAGTGTGAATTTGTAATGAAATTTTGGTGGAGC 328
Db 577 CCCATGTTCTGGCTTTCTTTTCAGCAAGCATGGCATAGTAATAAGAACTTTGGTGGAGC 636
Qy 329 GCTTTAGTTCAGGAGGTGCAGGTTCCAGAGGCTCGCTGTTTCTATGCTTTTCAAAATCTCA 388
Db 637 GATTTAGCAAGAGTTGAGATTTACAGAGCCCGCTGTTTCTATGGCTTTCCAAATTTGCCA 696
Qy 389 TCGAATATGTTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAAAACATACATTTCTGAATGTATAGTCTTCTTATTGACACTTTACATAAAGATCCCA 756
Qy 449 AGAAAGGGAATTTTATTAAATGCAATTTGAAACCATGCCCCTATGTTAAGAAAAAAGCAG 508
Db 757 AAGAAAGGGAATTTCTCTTCAATGCAATTTGAAACCATGCTTGTCTCAAGAAAGAGGCAG 816
Qy 509 ATTGGGCTTGGGATGGATAGACATAGAAATCTACTTTTGGGAAAGAGTGGTGGCT 568
Db 817 ACTGGGCTTGGCTGGATGGGACAAAGAGGCTACCTATGTTGTAACGCTGTTGTAGCCT 876
Qy 569 TTGCTGCTGTAGAGGAGTTTCTTCTCAGGATCTTTTGTCTGCTATATTTCTGGCTAAAAGA 628
Db 877 TTGCTGAGTGAAGGCAATTTTCTTTCCGGTCTTTTTCGGTCTGATATTTCTGGCTCAAGA 936
Qy 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCAATGAATCTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGAGTGTGCTGGCTCCTCACAATTTCTTAATGAACCTTATTAGCAGAGATGAGGTT 996
Qy 689 TTCACTGTGACTTTGCTTGCCTGATGTTCCAACTATTAGTAATAAGCCTTTCAGAGAAA 748
Db 997 TACACTGTGATTTTGGCTTGGCTGATGTTCAAAACACCTGGTACACAAACCATCGAGGAGA 1056
Qy 749 GGCTCAGGAGATCAATTTGTCATGCTGTCAAAATTTGAGCAGGAGTTTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATAATATCAATGCTGTTCCGATAGAACAGGAGTTTCTCACTGAGGCT 1116
Qy 809 TGCCAGTTGGCCTCAITTTGGAATGAATTTGATGAAACAGTACATTTGAGTTTGTAG 868
Db 1117 TGCCTGTGAAGCTCAITTTGGGATGAATTCGACTCTAATGAAGCAATACATTTGAGTTTGTG 1176
Qy 869 CTGACAGATTACTTTGTGGAATTTGATTTCTCAAAGGTTTTTTCAGGAGCAAGAAATCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAAGTGGGTTTTAGCAAGGTTTTTTCAGAGTAGAGAACCCATTTG 1236
Qy 929 ATTTTATGAAAAATTTCTTTAGAGGAAAAACAAATTTCTTTTACCTTGGATGCTG 988
```

Db 1237 ACTTTATGAGAATAATTTCTACTGGAAGGAAAGACTAACTTCTTTGAGAAAGAGAGTAGGCG 1296
Qy 989 AGTATCAGCGTTTTCAGTATGCGAGAAACACACAGATAAGCTCTTCCACTTGGATGCGAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGATGTCAGTCCAAACAGAGAATCTTTTACCTTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 7

US-09-962-677-9
; Sequence 9, Application US/09962677
; Patent No. 6759200
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE PHOSPHORYLASE GENE SEQUENCE
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING
; TITLE OF INVENTION: THE TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015003
; CURRENT APPLICATION NUMBER: US/09/962,677
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128, 1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc_feature
; LOCATION: 1636, 1738, 2259
; OTHER INFORMATION: n = c or t
US-09-962-677-9

Query Match 55.0%; Score 579; DB 4; Length 2500;

Best Local Similarity 75.0%; Pred. No. 6.4e-166;

Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

Qy 89 ATGAAGAGCACTCTTAAGAGAGTCTCGCGGTTTGTCTATCTTCCATCCAGTACC 148
Db 397 AGGATGAGCGCTGTGAGAGAAACCCCGCGCTTTGTCTATCTTCCCATCCAGTACC 456
Qy 149 CTGATATTTGGAATGTATAACAGCAGCAGGCTTCTCTCGACAGCAGAGAGTGC 208
Db 457 ATGATATCTGGCAGATGATTAAGAGCAGAGGCTTCTTTTGGACCGCGAGAGTTG 516
Qy 209 ACTTATCAAGGATCTCCCTCACTGGAACAAAGCTTAAGACAGATGAGAAGTACTTCATCT 268
Db 517 ACCTCTCNAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576

Qy 269 CTCACATCTTTAGCCTTTTTCAGCCAGTGTGAATGTAATGAAATTTTGGTGGAGC 328
Db 577 CCATGTTCTGGCTTTCTTTGACAGCAAGCGATGTCATAGTAAATGAAATTTTGGTGGAGC 636
Qy 329 GCTTTAGTCAGGAGGTCAGGTTCCAGAGGCTGCTGTTCTATGCGTTCCTCAATTTCTCA 388
Db 637 GATTTAGCCAAAGAGTTTCAGATTACAGAAGCCGCTGTTCTATGGCTTCCAAATTTGCCA 696
Qy 389 TCGAGAATGTTCACTCAGAGATGTACAGATTGCTGTGATAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAAACATACATTCTGAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 756
Qy 449 AGAAAAGGGAATTTTATTAATGCAATGAAACCATGCTCTATGTTAAGAAAAAGCAG 508
Db 757 AAGAAAAGGGAATTTCTCTCAATGCCATTGAAACGCTTGTGTCAAGAAGAGGCAG 816
Qy 509 ATTGGGCTTTCGATGATAGACAGATAAGAAATCTACTTTTGGGAAAGAGTGGTGGCT 568
Db 817 ACTGGGCTTTCGCTGATTTGGGACAAAGAGGCTACCTATGTTGAACGCTGTTGTAGCCT 876
Qy 569 TTGCTGCTGTAGAAGGAGTTTCTTCTCAGATCTTTTGTGCTGTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGCACTGGAAGGCAATTTCTTTCCGCTCTTTTGTGCTGATATTTCTGGCTCAAGA 936
Qy 629 AGAGAGGTCTTATGCCAGGACTCACTTTTCCAATGAATCACTCAGCAGAGATGAAGAC 688
Db 937 AACGAGGACTGATGCTGGCTTCACTTTCTAATGAATCTATTAGCAGAGATGAGGCTT 996
Qy 689 TTCACTGTGACTTTGCTTGGCTGTATGTTCCAATCTTGTAAATTAAGCCTTCAGAGAA 748
Db 997 TACACTGTGATTTTGTGCTGATGTTCAAAACCTGGTACACAAACCATCGAGGAGA 1056
Qy 749 GGGTCAGGAGATCATTTGTTGATGCTCTCAAAATTTGAGCAGAGGTTTAAACAGAGCT 808
Db 1057 GAGTAAGAGAAATTAATTAATCAATGCTGTTGGATAGAACAGAGTCTCTACTGAGGCT 1116
Qy 809 TGCAGTTGGCCTCATTTGGAATGAAATTTGATGAAACAGTACATTCAGTTTGTAG 868
Db 1117 TGCTGTGAAGCTCATTTGGGATGAATTCATCTTAATGAAGCAATACATTCAGTTTGTG 1176
Qy 869 CTGACAGATTACTTGTGGAATCTGGATCTCAAAAGTTTTCAGGACAGAAATCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACCTGGGTTTTCAGAGGTTTTCAGAGTAGAGAACCCATTG 1236
Qy 929 ATTTTATGGAAGCAATTTCTTGAAGGAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
Db 1237 ACTTTATGGAATATTTTCTACATGGAAGGAAAGACTAACTTCTTTGAGAGAGAGTAGCG 1296
Qy 989 AGTATCAGCGTTTTCAGTTATGTCAGAAACCAAGATAACGCTCTTCACTTGGATGAG 1048
Db 1297 AGTATCAGAGATGGGAGTGTGTCAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 8

US-09-949-016-2025

; Sequence 2025, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2025
 ; LENGTH: 2479
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-949-016-2025

Query Match 54.9%; Score 578.4; DB 4; Length 2479;
 Best Local Similarity 75.0%; Pred. No. 9.7e-166;
 Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY	89	ATGAAGACCACTCTTAAGAAAGAGTTCTCGCGGTTTGTCACTTTTCCAAATCCAGTACC	148
DB	397	AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTCACTTCCCATCGAGTACC	456
QY	149	CTGATATTTGGAAATGTATAACAGGCAAGGCTTCTTCTGGACAGCAGAGAGGTGC	208
DB	457	ATGATATCTGGCAGATGTATAAGAGCAGAGGCTTCTTTTGGACCGCGAGGAGTGG	516
QY	209	ACTTATCAAGGATCTCCCTCACTGGAACAGCTTAAGCAGATGAGAGTCTTCACT	268
DB	517	ACCTCTCAAGGACATTCAGCACTGGGAATCCCTGAAACCGGAGGAGATATTTATAT	576
QY	269	CTCACATCTTAGCTTTTTCAGCCAGATGATGAATGTAAATGAAATTTGGTGAGC	328
DB	577	CCATGTTCTGGCTTTCTTTCAGCAGCAGATGCGATGATGAATGAAATTTGGTGAGC	636
QY	329	GCTTATGTCAGAGGTCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTTCA	388
DB	637	GATTTAGCAAGAGTTTCAGATTACAGAGCCCGCTGTTTCTATGGCTTCAAATTTCCA	696
QY	389	TCGAGATGTTCACTCAGAGATGACAGTTGCTGTATGACACTTACATCAGAGATCCCA	448
DB	697	TGGAATACATACATCTGAAATGTATAGTCTTCTTATGACACTTACATGAAATGATCCCA	756
QY	449	AGAAAGGGAATTTTATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT	508
DB	757	AGAAAGGGAATTTTCTTCAATGCCATGAAACGATGCTTGTGTCAGAGAGGAGCAG	816
QY	509	ATTGGGCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	568
DB	817	ACTGGGCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	876
QY	569	TGCTCTCTGATGAGAGGTTTCTTCTCAGGATCTTTTGTGCTATATTTCTGCTAAGA	628
DB	877	TTGCTGCTGATGAGAGGTTTCTTCTTCCGCTTCTTTCGCTGATATTTCTGCTCAAGA	936
QY	629	AGAGAGTCTTATGCCAGGACTCACTTTTCCATGCAATGCAATGCAATGCAATGCAATGCAAT	688
DB	937	AACGAGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	996
QY	689	TTCACTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	748
DB	997	TACACTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1056
QY	749	GGGTGAGGAGATCATTTGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	808
DB	1057	GAGTAAGAGAAATATTAATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1116
QY	809	TGCCAGTTGGCTCATTTGAATGAATTTGCAATTTTGAATGAATTTGCAATTTTGAATGAAT	868
DB	1117	TGCTGTGAGCTCATTTGGGATGAATTTGCAATTTTGAATGAATTTGCAATTTTGAATGAAT	1176
QY	869	CTGACAGATTTACTTTGGAACTTGGATTTCTCAAGAGTTTTCAGGAGGAGTATTTTCAAGAGCCT	928
DB	1177	CAGACAGATTTACTTTGGAACTTGGATTTCTCAAGAGTTTTCAGGAGGAGTATTTTCAAGAGCCT	1236
QY	929	ATTTATGAAACATTTCTTTTAGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	988
DB	1237	ACTTTATGAGAAATTTCTTCACTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1296

QY 989 AGTATCAGCGTTTTCAGTATGAGGAGAAACACAGATTAACGTTTTCACCTTGGATGCG 1048
 DB 1297 AGTATCAGGAGTGGAGTGTCAAGTCCAAACAGAGAAATTTCTTTTACCTTGGATGCG 1356
 QY 1049 ATTT 1052
 DB 1357 ACTT 1360

RESULT 9
 US-09-023-655-1370
 ; Sequence 1370, Application US/09023655
 ; Patent No. 6607879
 ; GENERAL INFORMATION:
 ; APPLICANT: Cocks, Benjamin G.
 ; APPLICANT: Susan G. Stuart
 ; APPLICANT: Jeffrey J. Sellhauer
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
 ; NUMBER OF SEQUENCES: 1508
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 ; CURRENT APPLICATION DATA: US/09/023,655
 ; FILING DATE: HEREWITH
 ; CLASSIFICATION:
 ; PRIOR APPLICATION NUMBER:
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Zeller, Karen J.
 ; REGISTRATION NUMBER: 37,071
 ; REFERENCE/DOCKET NUMBER: PA-0001 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 1370:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2500 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GENE BANK
 ; CLONE: g36154
 ; US-09-023-655-1370

Query Match 54.9%; Score 578.4; DB 4; Length 2500;
 Best Local Similarity 75.0%; Pred. No. 9.7e-166;
 Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY	89	ATGAAGACCACTCTTAAGAAAGAGTTCTCGCGGTTTGTCACTTTTCCAAATCCAGTACC	148
DB	397	AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTCACTTCCCATCGAGTACC	456
QY	149	CTGATATTTGGAAATGTATAACAGGCAAGGCTTCTTCTGGACAGCAGAGAGGTGC	208
DB	457	ATGATATCTGGCAGATGTATAAGAGCAGAGGCTTCTTTTGGACCGCGAGGAGTGG	516
QY	209	ACTTATCAAGGATCTCCCTCACTGGAACAGCTTAAAGCAGATGAGAGTACTTTCATCT	268

Db 517 ACCTCTCCAGAGCAATTCAGCACTGGGAATCCTCGAAACCCGAGGAGAGATATTTATAT 576
Qy 269 CTCACATCTTACGCTTTTTCAGCGAGTGATGGAATGTAAATGAAAAATTTGGTGGAGC 328
Db 577 CCCATGTTCTGCTTCTTTCAGCAAGCATGGCATAGTAATGAAACTTTGGTGGAGC 636
Qy 329 GCTTTAGTCAGAGGTGCAAGTTCAGAGGCTCGCTGTTTCTATGGCTTTCAAAATCTCA 388
Db 637 GATTTAGCCAGAAAGTTCAGATTACAGAAAGCCGCTGTTTCTATGGCTTTCCAAATGGCA 696
Qy 389 TCGAAGATGTTTCACTCAGAGGTGACAGTTCCTGTGATAGACATTCATACAGATCCCA 448
Db 697 TGGAAAAACATATCTCGAAATGTATAGTCTTCTATTGACACTTACATAAAGATCCCA 756
Qy 449 AGAAAAGGAATTTTATTAATGCAATTAAGAAACCATGCTATGTTAAAGAAAAAGAGCAG 508
Db 757 AAGAAAGGAAATTTCTTCAATGCAATTAAGAAACCATGCTATGTTCAAGAAAGGAGCAG 816
Qy 509 ATTGGGCTTTCGATGGATGACAGATAGAAATCTACTTTTGGGAAAAAGAGTGGGCT 568
Db 817 ACTGGGCTTTCGCTGGAATGGGACAAAGAGGCTACCTATGTTGAAACGTTGTAGCCT 876
Qy 569 TTGCTGCTGTAGAGAGTTCCTCTCAGGATCTTTTCTGCTATATTTCTGCTTAAAGA 628
Db 877 TTGCTGCTGTAGAGAGTTCCTCTTCTTCCGCTTCTTTTCTGCTGATTTCTTCTCAGA 936
Qy 629 AGAGAGGCTTATGCCAGGACTCACATTTTTCAAATGAACTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCTGCGCTCACATTTCTTAATGAACCTTATAGCAGATGAGGTT 996
Qy 689 TTCACTGTGACTTTCCTGCTGATGTTCCAAATCTTAGTAATAAGCCTTCAGAAAGAA 748
Db 997 TACACTGTGATTTTCTGCTGATGTTCCAAACCTGTTACAAACCATCGGAGGAG 1056
Qy 749 GGGTCAGGAGATCAATTTGATGATGTTTCAAAATTCAGCAGGAGTTCCTCACTGAGGCT 928
Db 1177 CAGACAGACTTATGCTGGAAGTGGGTTTTCAGCAAGCTTTCAGAGTAGAGAACCCATTG 1236
Qy 929 ATTTATGAAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
Db 1237 ACTTTATGAGAAATATTTCACTGGAAGAAAGACTAACTTCTTTGAGAAAGAGTAGGCG 1296
Qy 989 AGTATCAGCGTTTTCAGTTATGGCAGAAACCAACAGATACGCTTTCACCTTGGATGCGAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGATGTCAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 10
US-09-949-016-145
; Sequence 145, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 145
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-145

Query Match 54.9%; Score 578.4; DB 4; Length 2500;
Best Local Similarity 75.0%; Pred. No. 9.7e-166;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

Qy 89 ATCAAGAGCCACTCTTAAGAAAGAGTTCGCGCGTGTGTCATCTTTTCCAATCCAGTACC 148
Db 397 AGATGAGCGCTGCTCGAGAGAAACCCGCGCTTGTGTCATCTTCCCATCGAGTACC 456
Qy 149 CTGATATTTGAAAAATGATATAACAGGCACAGGCTTCTCTTGACAGCAGAAAGAGTGG 208
Db 457 ATGATATCTGGCAGATGATATAAGAAAGCAGAGGCTTCTTTTGGACCGCCGAGGAGTTG 516
Qy 209 ACTTATCAAAAGGATCTCCCTCATCTGGAACAAGCTTTAAAGCAGATGAGAGTACTTCACT 268
Db 517 ACCTCTCAAGGACATTCAGCACTGGGAATCCCTGNAACCCGAGGAGAGATATTTTATAT 576
Qy 269 CTCACATCTTATAGCTTTTTCAGCCAGTGAATGTAATGTAATGTAATGTAATGTAATGTA 328
Db 577 CCATGTTCTGGCTTTCTTTGACAGCAGATGCAATGTAATGTAATGTAATGTAATGTAATG 636
Qy 329 GCTTTAGTCAGAGGAGTTCAGAGGCTTCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTTCA 388
Db 637 GATTTAGCCAAAGATTCAGATTACAGAGCCGCTGTTTCTATGGCTTTCCAAATTTGCCA 696
Qy 389 TCAGAGATGTTCTACTCAGAGATGACAGTTCGCTGATAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAAAACATACATCTCTGAAATGTATAGTCTTCTTATTGACACTTACATAAAGATCCCA 756
Qy 449 AGAAAGGGAATTTTATTTAAATGCAATGAAACCATGCGCTTATGTTAAAGAAAAAGCAG 508
Db 757 AAGAAAGGGAATTTCTCTTCAATGCCATTTGAAACGATGCTTGTCTCAAGAAAGAGGAG 816
Qy 509 ATTGGGCTTTCGATGATAGCAGATAGAAATCTACTTTTGGGAAAAAGAGTGGGCT 568
Db 817 ACTGGGCTTTCGCTGATTTGGGCAAAAGAGGCTACCTATGGTGAACGTTGTGTAGCCT 876
Qy 569 TTGCTGCTGTAGAGAGTTCCTTCTCAGAGATCTTTTGGCTGCTATATTTCTGGCTTAAAG 628
Db 877 TTGCTGCTGTAGAGAGTTCCTTCTTCTCGGCTTCTTTTTCGCTGCTGATATTTCTGGCTCA 936
Qy 629 AGAGAGTCTTATGCCAGGACTCACATTTTTCCAATGAACTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGACTGATGCTGCGCTCACATTTCTTAATGAACCTTATTAGCAGAGATGAGGTT 996
Qy 689 TTCACTGTGACTTTCGCTGCTGATGTTTCCAAATCTTAGTAATAAGCCTTCAGAAAGAAA 748
Db 997 TACACTGTGATTTTGTCTGCTGATGTTCCAAACCTGTTACAAACCATCGGAGGAG 1056
Qy 749 GGGTCAGGAGATCAATTTGATGATGTTTCAAAATTCAGCAGGAGTTCCTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATAATTAATCAATGCTGTTCCGATAGAACAGGAGTTCCTCACTGAGGCT 1116
Qy 809 TGCAGTTCGCTCATTTGGATGAAATGATGATGAAACAGTACATTCATGATTTGTAG 868
Db 1117 TGCCTGTGAAGCTCATTTGGGATGAAATGCACTTAAATGAACCAATGATGATTTGTGG 1176
Qy 869 CTGACAGATTAATTTGGAACCTTGGATCTTCAAGAGTTCCTCAAGGAGAAATTCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAAGTGGGTTTTCAGCAAGCTTTCAGAGTAGAGAACCCATTG 1236
Qy 929 ATTTATGAAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
Db 1237 ACTTTATGAGAAATATTTCACTGGAAGAAAGACTAACTTCTTTGAGAAAGAGTAGGCG 1296
Qy 989 AGTATCAGCGTTTTCAGTTATGGCAGAAACCAACAGATACGCTTTCACCTTGGATGCGAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGATGTCAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

Db 1237 ACTTATGAGATATTTCACTGGAAGGAAGACTAACTTCTTTTGAGAGAGTAGGCG 1296
 Qy 989 AGTATCAGCGTTTTCAGTATGCGAGAAACCAAGATACGTTTCCACCTTGATGACAG 1048
 Db 1297 AGTATCAGAGGATGGGAGTGATGTCAGTCCAAGTCCAAGAGAGATTTCTTTTACCTTGATGCTG 1356
 Qy 1049 ATTT 1052
 Db 1357 ACTT 1360

RESULT 11
 US-08-905-223-125
 ; Sequence 125, Application US/08905223
 ; Patent No. 6222029
 ; GENERAL INFORMATION:
 ; APPLICANT: Edwards, Jean-Baptiste D.
 ; APPLICANT: Duclert, Aymeric
 ; APPLICANT: Lacroix, Bruno
 ; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
 ; NUMBER OF SEQUENCES: 503
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Knobbe, Martens, Olson & Bear
 ; STREET: 501 West Broadway
 ; CITY: San Diego
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 92101-3505
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy Disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: Win95
 ; SOFTWARE: Word
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/905,223
 ; FILING DATE:
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Israel, Ned A.
 ; REGISTRATION NUMBER: 29,655
 ; REFERENCE/DOCKET NUMBER:
 ; TELEPHONE: (619) 235-8550
 ; TELEFAX: (619) 235-0176
 ; INFORMATION FOR SEQ ID NO: 125:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 481 base pairs
 ; TYPE: NUCLEIC ACID
 ; STRANDEDNESS: DOUBLE
 ; TOPOLOGY: LINEAR
 ; MOLECULE TYPE: CDNA
 ; ORIGINAL SOURCE:
 ; ORGANISM: Homo Sapiens
 ; TISSUE TYPE: Brain
 ; FEATURE:
 ; NAME/KEY: sig peptide
 ; LOCATION: 41..343
 ; IDENTIFICATION METHOD: Von Heijne matrix
 ; OTHER INFORMATION: score 4.4
 ; OTHER INFORMATION: seq ISHILAFPAASDG/IV

US-08-905-223-125
 Query Match 41.9%; Score 441; DB 3; Length 481;
 Best Local Similarity 100.0%; Pred. No. 3.2e-124;
 Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGGCGCCCGAAGCGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
 Db 41 ATGGGCGCCCGAAGCGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA 100
 Qy 61 GACACCAACGAAGTGAATTAAGTCAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGC 120
 Db 101 GACACCAACGAAGTGAATTAAGTCAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGC 160

Qy 121 CGGTTTGTCTCTTCCATCCAGTACCTGATATTTGGAAATGTATATAACAGGCACAG 180
 Db 161 CGGTTTGTCTCTTCCATCCAGTACCTGATATTTGGAAATGTATATAACAGGCACAG 220
 Qy 181 GCTTCCCTTCTGGACAGAGAGGTCGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 240
 Db 221 GCTTCCCTTCTGGACAGAGAGGTCGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 280
 Qy 241 CTTAAAGCAGATGAGAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCGAGCCAGTGAT 300
 Db 281 CTTAAAGCAGATGAGAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCGAGCCAGTGAT 340
 Qy 301 GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGAGGTCAGGTCAGGTCAGAGGCT 360
 Db 341 GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGAGGTCAGGTCAGAGGCT 400
 Qy 361 CGCTGTTTCTATGCTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
 Db 401 CGCTGTTTCTATGCTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 460
 Qy 421 CTGATAGACACTTACATCAGA 441
 Db 461 CTGATAGACACTTACATCAGA 481

RESULT 12
 US-08-307-499-1/c
 ; Sequence 1, Application US/08307499
 ; Patent No. 5651972
 ; GENERAL INFORMATION:
 ; APPLICANT: Moyer, Richard W.
 ; APPLICANT: Vi uela, Eladio
 ; APPLICANT: Gibbs, E.P.J.
 ; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
 ; NUMBER OF SEQUENCES: 60
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: David R. Saliwanchik
 ; STREET: 2421 N.W. 41st Street, Suite A-1
 ; CITY: Gainesville
 ; STATE: Florida
 ; COUNTRY: U.S.A.
 ; ZIP: 32606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/307,499
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/908,241
 ; FILING DATE: 1-JUL-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/908,630
 ; FILING DATE: 29-JUN-1992
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/342,212
 ; FILING DATE: 21-APR-1992
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Saliwanchik, David R.
 ; REGISTRATION NUMBER: 31,794
 ; REFERENCE/DOCKET NUMBER: UP35.1.FWCCI
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 904-375-8100
 ; TELEFAX: 904-372-5800
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:


```

; LENGTH: 14176 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 3852..4226
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 4585..4887
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 5131..5310
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 5760..5912
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 6786..7130
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 10148..10513
;   LOCATION: 10148..499-1
;
Query Match
Best Local Similarity 40.8%; Score 429.8; DB 1; Length 14176;
Matches 651; Conservative 0; Mismatches 327; Indels 5; Gaps 2;

QY 71 AAAGTGAATAAGTCAAAATGAAGACCCACTCTCTAAGAAAGAGTTCGCGCGGTTGTCA 130
DB 10841 AAAATGATTTTATACACGAATGGAGCCCTATTCTTCAAGAGTCTGATTTCTAGGTTGGTTA 10782

QY 131 TCTTTCCATCCAGTACCTCTGATATTGGAAATGTATAAACAGGACACAGGCTTCTTCT 190
DB 10781 TTTTCCCTATTAAAGTATCATGATATCTGGAATATGTATAAACAATCAGTGGCAAGTTTT 10722

QY 191 GGACGACGAAGAGGTGCACTTATCAAAAGATCTCCTCACTGGAACAAGCTTAAAGCAG 250
DB 10721 GGACCGTGAAGAGTAGATATTATCAAAAGATTTAGATGATGGGATTAATTAACCTAAG 10662

QY 251 ATGAGAAGTACTTCACTCTCTCAATCTTAGCCCTTTTTCGAGCCAGTGATGGAATTGTAA 310
DB 10661 ACGAAAATACITTTATAAAACATATACTAGCAITTTTTCGATCTAGTGATGGTATTGTAA 10602

QY 311 ATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTTCAGAGGCTCGCTGTTTCT 370
DB 10601 ATGAGAATTTAGCGGAAGATTTTATGTGGATGTACAGTGTTCAGAGGCCAGATGTTCT 10542

QY 371 ATGGCTTTCAAAATCTCATCGAATGTTTCACCTCAGAGATGTACAGTTGCTGTAGACA 430
DB 10541 ATGGATTTCAAAATAGCTATGGAAAATTTTCATTCAGAAATGTATAGTTTATTAATAGATA 10482

QY 431 CTTATCATCAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGAAACCATGCGCT 490
DB 10481 CATATCTAGAGATATATAGAAAATATGCAITTTTAAACGCTATAGAAACATGGAAT 10422

QY 491 ATGTTAAGAAAAAAGCAGATTTGGGCTTTCGATGGATAGCAGATAGAAAATCTACTTTTG 550
DB 10421 GCGTAAAAAAGAAAGCTGATTTGGGCCAAGAAATGGATATC---TAGCAACAAGGTATATG 10365

QY 551 GGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAAGGATTTTCTTCTCAGGATCTTTTGCTG 610
DB 10364 GAGAAAGAGTAGTAGCATTTTGCAGCTTGGAGGGAATTTCTTTCTGGTTCAATTTGCTG 10305

QY 611 CTATATCTGGCTTAAAGACAGAGGTTCTTATGCGAGGACTCACTTTTCCAAATGAACTCA 670
DB 10304 CTATATTTGGATTAATAAAGACAGAGGATGTATGCGGATTAACATTTTCTTAAGAACTAA 10245

QY 671 TCAGCAGATGTAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTTCCAAATCTTAGTAA 730
DB 10244 TAAGTAGACGAAAGGTTTACATTTGATTTTGGTGTGTTTAAATGTTTAAACAATTTATAC 10185

```

```

QY 731 ATAAAGCTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTCTCAAAATTCAGCAGG 790
DB 10184 ATCCACCATCTAAGGAAGTTTATAACGTCGATAATCATTTGATGCGGTTAATATAGAAAAG 10125

QY 791 AGTTTTTAACAGAAAGCCTTCGCCAGTTGCGCTCATTTGGAATGAAATTCGATTTTGAATAAC 850
DB 10124 AGTTTTTGACAGTTGCTATTTCGGTGGATCTTATAGGTATGAATTTGTTTAAATGCTC 10065

QY 851 AGTACATTTGAGTTTGTAGCTGCACAGATTACTTGTGGAACTTTGGATTCTCAAAGGTTTTTC 910
DB 10064 AGTATATAGAAATTCGTCGCAGATAGATTATTAAACAGAGTTAGGTTGTCAAAG--TCTCA 10007

QY 911 AGGCAAGAAATCCTTTTGTATTTATGGAACATTTCTTTAGAGAGAAAAACAATTTCT 970
DB 10006 ATGTATATATCCTTTTGTAGCTTTTATGAGTATATATCATAGAAAGGTAAGACTAATTTT 9947

QY 971 TTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATTATGCGAGAAACACAGATAAG 1030
DB 9946 TCGAAACGACGAGTTAGTGAATATCAAAAGATGGGGGTGTTTACAAATTAAGAAAGATA 9887

QY 1031 TCTTCACTTGGATGCAGATTTT 1053
DB 9886 TATTTCTACGGATATAGATTTT 9864

RESULT 13
US-08-307-499-14
; Sequence 14, Application US/08307499
; Patent No. 5651972
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,499
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF35.1.FWCCI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14176 base pairs

```

```
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 138..1460
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 2456..2659
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 2809..3030
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 3070..3330
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 3356..4180
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 4392..5894
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 6171..6398
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 6447..6875
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 6928..7431
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 7454..7858
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 7895..8155
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 8215..8682
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 8715..9539
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 9562..10272
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 10316..11908
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 11971..12780
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 12829..13107
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 13149..14171
; US-08-307-499-14
```

```
Query Match 40.8%; Score 429.8; DB 1; Length 14176;
Best Local Similarity 66.2%; Pred. No. 5.9e-120;
Matches 651; Conservative 0; Mismatches 327; Indels 5; Gaps 2;

QY 71 AAGTGAATTAAGTCAAGAGCCACTCTTAAGAAAGAGTCTCGCGGTTGTCA 130
    |||
Db 3336 AAAATGATTTTATACAGAAATGAGGCTATTCTTCAAGAGTCTGATTTCTAGGTTGCTTA 3395

QY 131 TCTTTCATCCAGTACCTGATATTGGAAATGTATAACAGGACAGGCTTCTTCT 190
    |||
Db 3396 TTTTCCCTATTAGTATCATGTATCTGAAATATGTATAACCAATCAGTGGCAAGTTT 3455

QY 191 GGACAGCAGAGAGGTGCGACTTATCAAGAGTCTCCCTCACTCGAACAAGCTTTAAAGCAG 250
```

```
Db 3456 GGACGTTGAAGAGTAGATTATCAAAAGATTTATAGATGATTGGGATAAATTAACCTAAG 3515
QY 251 ATGAGAAGTACTTCTATCTCTCACATCTTTAGCCCTTTTTCGACCCAGTGTATGGAATGTAA 310
    |||
Db 3516 ACGAAAATACTTTATATAAACAATATCTAGCATTTTTCATCTAGTGTATGTAATGTAA 3575
    |||
QY 311 ATGAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGTCGCTGTTTCT 370
    |||
Db 3576 ATGAGATTAGCGGAAAGATTTTATGTGGATGTACAGTGTTCAGAGGACGATGTTTCT 3635
    |||
QY 371 ATGGCTTTCAATTTCTCATCGAGATGTTCTACAGAGATGTACAGTGTTCCTGATAGACA 430
    |||
Db 3636 ATGGATTTCAAATAGCTATGGAATAATTCATTCAGAAATGTATAGTTTATTAATAGATA 3695
    |||
QY 431 CTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTTGAAACCATGCCCT 490
    |||
Db 3696 CATATGTAAGAGATAATATAGAAAATATGCAATTTATTAACGCTATAGAAAACAATGGAAT 3755
    |||
QY 491 ATGTTAAGAAAAGACAGATTGGGCTTTGCGATGATAGCAGATAGAAAATCTACTTTTG 550
    |||
Db 3756 GCGTAAAAAAGAAAGCTGATTGGGCCAGAAAATGGAATC---TAGCAACAAGGTATATG 3812
    |||
QY 551 GGGAAAGAGTGTGCGCTTTGCTGCTGTAGAGAGTTCCTCTCAGGATCTTTTGCTG 610
    |||
Db 3813 GAGAAAGAGTAGTAGCATTTTCAGCTGTGGAGGGAATATCTTTTCTGGTTCAATTTGCTG 3872
    |||
QY 611 CTATATTCTGGCTAAAGAGAGAGCTTTATGCGCAGGACTCACTTTTCCAAATGAACTCA 670
    |||
Db 3873 CTATATTTTGGATTAATAAACAAGAGATTGATGCGCGGATTAACATTTCTAATGAACATA 3932
    |||
QY 671 TCAGCAGAGATGAAGGACTTCACCTGTGACTTTTGTCTTGCTGCTGATGTTCCAATCTTAGTAA 730
    |||
Db 3933 TAAGTAGAGACGAAGTTTACATTGTGATTTTGGGTGTTTAAATGTTTAAACATTTATTAC 3992
    |||
QY 731 ATAGGCTTTCAGAAAGAGGTGAGGAGATCATTTGTTGATGCTCTCAAAATTTGACGAG 790
    |||
Db 3993 ATCCACCATCTTAAGGAAGTTATAACGTCGATAATCATTTGATGCGGTAAATATAGAAAAG 4052
    |||
QY 791 AGTTTTTAACAGAGCCCTTGCCAGTTGGCCTCATTTGGAATGAATGCAATTTTGTGATCAAC 850
    |||
Db 4053 AGTTTTTGACAGTTGCTATTCCGGTGGATCTTAGGTTGATGAATTTGTTGTTTAAATGCTC 4112
    |||
QY 851 AGTACATTGAGTTGTAGTCAGACATTAATCTGTGGAACCTTGGATCTCAAAAGTTTTC 910
    |||
Db 4113 AGTATATAGAAATTCGTCAGATAGATTAATAACAGATTAGTTGTGAAAAG--TCTCA 4170
    |||
QY 911 AGCAGAAAATCTTTTGAATTTTATGGAACATTTCTTTAGAGGAAAAACAATTTTCT 970
    |||
Db 4171 ATGTATATAATCTTTTAGCTTTTATGGAGTATATACACTAGAAGGTAAGACTAATTTT 4230
    |||
QY 971 TTGAAAACGAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACCAAGATAACG 1030
    |||
Db 4231 TCGAACGACGAGTTAGTGAATATCAAAAGATGGGGGTGTTTACAAATAAAGAGAAATA 4290
    |||
QY 1031 TCTTCACCTTGGATGCGAGATTTT 1053
    |||
Db 4291 TATTTCTACGATATAGATTTT 4313
    |||
```

```
RESULT 14
US-09-299-268-1/c
; Sequence 1, Application US/09299268
; Patent No. 6217882
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, B.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
```

STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.
ZIP: 32606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,268
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/901,127
FILING DATE:
APPLICATION NUMBER: US 07/908,241
FILING DATE: 1-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,630
FILING DATE: 29-JUN-1992
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/342,212
FILING DATE: 21-APR-1992
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 14176 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)

FEATURE:
NAME/KEY: CDS
LOCATION: 3852..4226
FEATURE:
NAME/KEY: CDS
LOCATION: 4585..4887
FEATURE:
NAME/KEY: CDS
LOCATION: 5131..5310
FEATURE:
NAME/KEY: CDS
LOCATION: 5760..5912
FEATURE:
NAME/KEY: CDS
LOCATION: 6786..7130
FEATURE:
NAME/KEY: CDS
LOCATION: 10148..10513

Query Match 40.8%; Score 429.8; DB 3; Length 14176;
Best Local Similarity 66.2%; Pred. No. 5.9e-120;
Matches 651; Conservative 0; Mismatches 327; Indels 5; Gaps 2;

QY 71 AAAGTGAATAAGTCAAGAGAGAGTCTCTAAGAGAGAGTCTCGCCGGTTGTCA 130
DB 10841 AAAATGATTTTATACAGAAATGAGGCTATTCTTCAAGAGTCTGATCTAGTTCGTTA 10782
QY 131 TCTTTCCAAATCCAGTACCTTGATTTTGGAAATGTATAAACAAGGCAAGGCTTCCTTCT 190
DB 10781 TTTTCCCTATTAAAGTATCATGATATCTGGAAATGTATAAACAATCATGAGTGGCAAGTTT 10722

QY 191 GGACAGAGAGAGAGTGGAGTCTATCAAGAGATCTCCCTCACTGGAACAGCTTAAAGCAG 250
DB 10721 GGACCGTGAAGAGTAGATTTATCAAAAGATTTAGATGATTTGGGATAAATTAACATAAG 10662
QY 251 ATGAGAGTACTTCACTCTCAATCTAGCCTTTTGGAGCCAGTCAATGGAATTTGTA 310
DB 10661 ACAGAAATATCTTTATTAACATATATAGCAATTTTGGCATCTAGTGATGATTTGTA 10602
QY 311 ATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTTCCAGAGGCTCGCTGTTCT 370
DB 10601 ATGAGATTTAGCGGAAGATTTTATGTGGATGTACAGTGTTCAGAGGCAAGATTTCT 10542
QY 371 ATGCTTTCAATTTCTCATCGAGATGTTCTCAGAGATGTACAGTGTTCAGTGTAGACA 430
DB 10541 ATGGAATTTCAATAGCTATGGAATAATTCATTGAGAAATGATAGTTTATTAATAGATA 10482
QY 431 CTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTCAGAACCATGCCT 490
DB 10481 CATATGTAAGAGATAATATAGAAAAATGCAATTTATTTAACGCTATAGAAAACATGGAAT 10422
QY 491 ATGTTAAGAAAAAGCAGATTTGGGCTTGGCATGATAGCAGATAGAGATAAGAAATCTACTTTG 550
DB 10421 GCGTAAAAAGAAAGCTGATTTGGCCAGAAAAATGATATC--TAGCAACAAGGTATATG 10365
QY 551 GGGAAAGAGTGGTGGCCTTTTCTGTCTGTAGAGAGTGTTCCTCTCAGGATCTTTTGTG 610
DB 10364 GAGAAAGAGTAGTAGCATTTGCACTGTGGAGGGAATATCTTTTCTGTTCAITTTGCTG 10305
QY 611 CTATATTCGCTAAGAGAGAGTCTTATGCGAGGATTCACATTTTCCATGAACTCA 670
DB 10304 CTATATTTTGGATAAAAAACGAGGATTTGATGCGCGGATTAACATTTTCTAATGAACTAA 10245
QY 671 TCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTCATGTTCCAATCTACTAGTAA 730
DB 10244 TAAGTAGAGAGAGGTTTACATTTGCTGTGTTAATGTTTAAACATTTATTATAC 10185
QY 731 ATAAGCCTTCAGAAAGAGGTTCAGGAGATCATTTGTTGATGCTGTCAAAATTTGAGCAGG 790
DB 10184 ATCCACCATCTAAGGAAGTTTATACGTCGATATCAITTCATGTCGTTAATATAGAAAAG 10125
QY 791 AGTTTTTAACAGAGCCTTGCAGTTCGCTCATTTGGAATGCAATTTGATGAAAC 850
DB 10124 AGTTTTTGCAGTGTCTTATTCGGTGGATCTTATAGGTATGAATTTGTTTAAATGCTC 10065
QY 851 AGTACATTTGATTTGTAGCTGACAGATTACTTTGGAACTTGGATTTCTCAAGGTTTTTC 910
DB 10064 AGTATATAGATTCGTCGAGATAGATTTTAAACAGAGTTAGTTTGTGAAAAG--TCTCA 10007
QY 911 AGCAGAAAAATCCTTTTGTATTTTATGAAAAATTTCTTTTGAAGAAAAACAAATTTCT 970
DB 10006 ATGTATATATCCTTTTATGCTTTTATGGAGTATATATCACTAGAAGGTAAGACTAATTTT 9947
QY 971 TTGAGAAACAGTTTCAGAGTATCAGGCTTTTGGAGTTATGCGAGAAACCAAGATAAG 1030
DB 9946 TCGAACGACGAGTTAGTGAATATCAAAAGATGCGGGGTGTTTACAAATAAAGAGAAATA 9887
QY 1031 TCTTCACCTTTGGATGCAATTTT 1053
DB 9886 TATTTTCTACGATATAGATTTT 9864

RESULT 15
US-09-299-268-14
; Sequence 14, Application US/09299268
; Patent No. 6217882
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:

Qy	731	ATAAGCCTTCAGAGAAAGGTCAGGAGATCATTTGATGCTGCAAAATTTGAGCAGG	790
Db	3993	ATCCACCATCTAAGGAAGTTATAACGTCGATAATCATTCGATCGGTTAATATAGAAAAGG	4052
Qy	791	AGTTTTTAACAGAACCTTGCAGTTGGCCTCATTTGGAATGCAATTTTGATGAAC	850
Db	4053	AGTTTTTGACAGTTGCTATTCGGTGGATCTTATAGGTATGAATTTGTTTAATGTCTC	4112
Qy	851	AGTACATTGAGTTTGTAGCTGCAGAGATTACTTGTGGAACTTGGATTCTCAAAGGTTTTTC	910
Db	4113	AGTATATAGAAATTCGTCGAGATAGATTATTAACAGAGTTAGTTGTTGAAAAG--TCTCA	4170
Qy	911	AGGCAGAAATCCTTTTGTATTTATGGAACAAATTTCTTTAGAAAGGAAAAACAAATTTCT	970
Db	4171	ATGTATATAATCCTTTTAGCTTTATGGAGTATATATCACTAGAAAGGTAAGACTAAATTTT	4230
Qy	971	TTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACCAAGATAACG	1030
Db	4231	TCGAACGACGAGTTAGTGAATATCAAAAGATGGGGGTGTTTACAAATAAAGNAGAGAATA	4290
Qy	1031	TCCTTACCTTGGATGCAGATTTT	1053
Db	4291	TATTTCTACGGATATAGATTTT	4313

Search completed: June 16, 2005, 09:43:26
Job time : 224 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2005, 08:31:43 ; Search time 753 Seconds
(without alignments)
8680.672 Million cell updates/sec

Title: US-10-698-228-2
Perfect score: 1053
Sequence: 1 atggggcaccggaaagcc.....tcacattggtgacagatttt 1053

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 6054689 seqs, 3103772919 residues

Total number of hits satisfying chosen parameters: 12109378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications, NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/2/pubpna/US10J_PUBCOMB.seq:*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1053	100.0	1053	18	US-10-698-228-2
2	1053	100.0	4955	17	US-10-172-118-71
3	1053	100.0	4955	18	US-10-342-887-71
4	1053	100.0	4955	18	US-10-698-228-3
5	1051.4	99.8	1053	18	US-10-698-228-12
6	1051.4	99.8	1081	18	US-10-698-228-4
7	578.4	54.9	1989	9	US-09-925-301-505
					Sequence 2, Appli
					Sequence 71, Appl
					Sequence 3, Appli
					Sequence 12, Appl
					Sequence 4, Appli
					Sequence 505, App

8	578.4	54.9	2216	15	US-10-084-817-342	Sequence 342, App
9	578.4	54.9	2500	9	US-09-954-456-724	Sequence 724, App
10	578.4	54.9	2500	9	US-09-954-456-1169	Sequence 1169, App
11	578.4	54.9	2500	9	US-09-954-456-1827	Sequence 1827, App
12	578.4	54.9	2500	18	US-10-641-643-1370	Sequence 1370, Ap
13	578.4	54.9	2500	20	US-10-733-878-458	Sequence 458, App
14	578.4	54.9	2500	21	US-10-843-641A-3751	Sequence 3751, Ap
15	578.4	54.9	2500	21	US-10-843-641A-4136	Sequence 4136, Ap
16	578.4	54.9	2500	21	US-10-843-641A-4854	Sequence 4854, Ap
17	578.4	54.2	2113	21	US-10-764-420-1636	Sequence 1636, Ap
18	561	53.3	1328	18	US-10-403-571-75	Sequence 75, Appl
19	545.4	51.8	977	17	US-10-264-237-790	Sequence 790, App
20	447.8	42.5	186854	21	US-10-872-156-34	Sequence 34, Appl
C	447.8	42.0	1146	9	US-09-822-830A-174	Sequence 174, App
21	432.2	41.0	1218	15	US-10-128-714-7245	Sequence 7245, Ap
22	432.2	38.3	963	17	US-10-369-493-34052	Sequence 34052, A
23	403.2	38.3	963	17	US-10-369-493-34052	Sequence 2245, Ap
24	397	37.7	1146	15	US-10-128-714-2245	Sequence 46252, A
25	392.6	37.3	1173	17	US-10-369-493-46252	Sequence 6537, A
26	392	37.2	1242	16	US-10-032-585-6537	Sequence 27809, A
27	386.6	36.7	1292	17	US-10-369-493-27809	Sequence 126843,
28	380.6	36.1	1248	18	US-10-424-599-126843	Sequence 45857, A
29	380.4	36.1	1200	17	US-10-369-493-45857	Sequence 40471, A
30	371.2	35.3	1361	18	US-10-424-599-40471	Sequence 73024, A
31	365.8	34.7	1586	18	US-10-424-599-73024	Sequence 2145, Ap
32	364.8	34.6	1206	17	US-10-320-797-2145	Sequence 43, Appl
33	361.6	34.3	957	20	US-10-653-047-43	Sequence 1245, Ap
34	349	33.1	1314	15	US-10-128-714-1245	Sequence 6245, Ap
35	349	33.1	1450	15	US-10-128-714-6245	Sequence 245, App
36	349	33.1	3314	15	US-10-128-714-245	Sequence 5245, Ap
37	349	33.1	3450	15	US-10-128-714-5245	Sequence 12381, A
38	328.2	31.2	560	9	US-09-864-761-12381	Sequence 31361, A
39	324	30.8	1169	18	US-10-425-114-31361	Sequence 31182, A
40	324	30.8	2992	20	US-10-425-115-31182	Sequence 15462, A
41	318.6	30.3	1569	20	US-10-425-115-15462	Sequence 1378, Ap
42	314.4	29.9	1306	20	US-10-739-930-1378	Sequence 13723, A
43	312.8	29.7	1255	19	US-10-767-701-13723	Sequence 63321, A
44	309.6	29.4	1791	20	US-10-425-115-63321	Sequence 11109, A
C	306.2	29.1	1725	19	US-10-437-963-11109	

ALIGNMENTS

RESULT 1

US-10-698-228-2
; Sequence 2, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
; US-10-698-228-2

Query Match 100.0%; Score 1053; DB 18; Length 1053;
Best Local Similarity 100.0%; Pred. No. 7.9e-295;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

QY 1 ATGGGCGACCCGGAAGGCGGGAAGGCGGCGGCTGGATCAGATGAGATGATCATCTTCA 60
Db |||||
QY 1 ATGGGCGACCCGGAAGGCGGGAAGGCGGCGGCTGGATCAGATGAGATGATCATCTTCA 60
Db |||||
QY 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGGCCACTCTTAAGAAAGTTCCTCGC 120
Db |||||
QY 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGGCCACTCTTAAGAAAGTTCCTCGC 120
Db |||||
QY 121 CGGTTTGTCTATCTTTCAATCCAGTACCTGATATTTGGAAATGATATAACAGGCACAG 180
Db |||||
QY 121 CGGTTTGTCTATCTTTCAATCCAGTACCTGATATTTGGAAATGATATAACAGGCACAG 180
Db |||||
QY 181 GCTTCCTCTTGGACAGACAGAGGTCGACTTATCAAGAGATCTCCCTCACTGSAACAAG 240
Db |||||
QY 181 GCTTCCTCTTGGACAGACAGAGGTCGACTTATCAAGAGATCTCCCTCACTGSAACAAG 240
Db |||||
QY 241 CTTAAAGCAGATGAGAGTACTTCACTCTCACTTACGCTTTTGGAGCCAGTGAT 300
Db |||||
QY 241 CTTAAAGCAGATGAGAGTACTTCACTCTCACTTACGCTTTTGGAGCCAGTGAT 300
Db |||||
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAAGTTTCCAGAGCT 360
Db |||||
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAAGTTTCCAGAGCT 360
Db |||||
QY 361 CGCTGTTTCTATCGCTTTCAATTTCTCATCGAGAACTTTCACTCAGAGATGTACAGTTG 420
Db |||||
QY 361 CGCTGTTTCTATCGCTTTCAATTTCTCATCGAGAACTTTCACTCAGAGATGTACAGTTG 420
Db |||||
QY 421 CTGATAGACACTTACATCAGAGATGCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 480
Db |||||
QY 421 CTGATAGACACTTACATCAGAGATGCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 480
Db |||||
QY 481 ACCATGCCCTATGTTAAGAAAGGAGATGAGGCTTTGGAGTGGATGATAGAGAA 540
Db |||||
QY 481 ACCATGCCCTATGTTAAGAAAGGAGATGAGGCTTTGGAGTGGATGATAGAGAA 540
Db |||||
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTAGAGAGGATTTTCTTCTCAGGA 600
Db |||||
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTAGAGAGGATTTTCTTCTCAGGA 600
Db |||||
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGGCTTTATGCCAGGACTCACTTTTTC 660
Db |||||
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGGCTTTATGCCAGGACTCACTTTTTC 660
Db |||||
QY 661 AATGAACATCATCAGAGATGAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db |||||
QY 661 AATGAACATCATCAGAGATGAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db |||||
QY 721 TACTTAGTAAATGAAGCTTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db |||||
QY 721 TACTTAGTAAATGAAGCTTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db |||||
QY 781 ATTGACAGAGTATTTTAAAGAGGCTTCCAGTGGCTTCACTTGAATGAATTCATT 840
Db |||||
QY 781 ATTGACAGAGTATTTTAAAGAGGCTTCCAGTGGCTTCACTTGAATGAATTCATT 840
Db |||||
QY 841 TTGATGAACAGTACATTTGATTTGTAGCTGACAGATTAATTGTTGAACTTGGATTCTCA 900
Db |||||
QY 841 TTGATGAACAGTACATTTGATTTGTAGCTGACAGATTAATTGTTGAACTTGGATTCTCA 900
Db |||||
QY 901 AAGGTTTTTTCAGGAGAAATCTTTTGAATTTTGAATTTTGAATTTTGAAGGAA 960
Db |||||
QY 901 AAGGTTTTTTCAGGAGAAATCTTTTGAATTTTGAATTTTGAATTTTGAAGGAA 960
Db |||||
QY 961 ACAAAATTTTTCAGAAACAGTTTTCAGAGTATCAGGCTTTTTCAGTTATGCGAGAAACC 1020
Db |||||
QY 961 ACAAAATTTTTCAGAAACAGTTTTCAGAGTATCAGGCTTTTTCAGTTATGCGAGAAACC 1020
Db |||||
QY 1021 ACAGATAACGCTTTCACCTTGGATGCAATTTT 1053
Db |||||
QY 1021 ACAGATAACGCTTTCACCTTGGATGCAATTTT 1053
Db |||||

```

```

RESULT 2
US-10-172-118-71
; Sequence 71, Application US/10172118
; Publication No.: US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AB036063
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-71

```

```

Query Match 100.0%; Score 1053; DB 17; Length 4955;
Best Local Similarity 100.0%; Pred. No. 2e-294;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGGCGGGAAGGCGGCGGCTGGATCAGATGAGATGATCATCTTCA 60
Db |||||
QY 245 ATGGGCGACCCGGAAGGCGGGAAGGCGGCGGCTGGATCAGATGAGATGATCATCTTCA 304
Db |||||
QY 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGGCCACTCTTAAGAAAGTTCCTCGC 120
Db |||||
QY 305 GACACCAACGAAAGTGAATTAAGTCAATGAAGGCCACTCTTAAGAAAGTTCCTCGC 364
Db |||||
QY 121 CGGTTTGTCTATCTTTCAATCCAGTACCTGATATTTGGAAATGATATAACAGGCACAG 180
Db |||||
QY 365 CGGTTTGTCTATCTTTCAATCCAGTACCTGATATTTGGAAATGATATAACAGGCACAG 424
Db |||||
QY 181 GCTTCCTCTTGGACAGACAGAGGTCGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db |||||
QY 425 GCTTCCTCTTGGACAGACAGAGGTCGACTTATCAAGAGATCTCCCTCACTGGAACAAG 484
Db |||||
QY 241 CTTAAAGCAGATGAGAGTACTTCACTCTCAGATCTTTAGCCTTTTTCAGCCAGTGAT 300
Db |||||
QY 485 CTTAAAGCAGATGAGAGTACTTCACTCTCAGATCTTTAGCCTTTTTCAGCCAGTGAT 544
Db |||||
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGGTTCCAGAGGCT 360
Db |||||
QY 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGGTTCCAGAGGCT 604
Db |||||
QY 361 CGCTGTTTCTATGCTTTCAAAATTTCTATCAGAGATGTTTCACTCAGAGATGTACAGTTG 420
Db |||||
QY 605 CGCTGTTTCTATGCTTTCAAAATTTCTATCAGAGATGTTTCACTCAGAGATGTACAGTTG 664
Db |||||
QY 421 CTGATAGACACTTACATCAGAGATGCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 480
Db |||||
QY 665 CTGATAGACACTTACATCAGAGATGCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 724
Db |||||
QY 481 ACCATGCCCTATGTTAAGAAAGGAGATGAGGCTTTGGAGTGGATGATAGAGAA 540
Db |||||
QY 725 ACCATGCCCTATGTTAAGAAAGGAGATGAGGCTTTGGAGTGGATGATAGAGAA 784
Db |||||
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTAGAGAGGATTTTCTTCTCAGGA 600
Db |||||
QY 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTAGAGAGGATTTTCTTCTCAGGA 844
Db |||||

```

QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGAGGCTTTATAGCAGAGCTACATTTTTC 660
 DB 845 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGAGGCTTTATAGCAGAGCTACATTTTTC 904
 QY 661 AATGAACCTCATCAGCAGAGATGAAGGCTTCACTGTGACATTTGCTTGCCTGATGTTCCAA 720
 DB 905 AATGAACCTCATCAGCAGAGATGAAGGCTTCACTGTGACATTTGCTTGCCTGATGTTCCAA 964
 QY 721 TACTTAGTAAATAAGGCTTCAAGAGAGAGAGGCTTCACTGTGACATTTGCTTGCCTGATGTTCCAA 780
 DB 965 TACTTAGTAAATAAGGCTTCAAGAGAGAGAGGCTTCACTGTGACATTTGCTTGCCTGATGTTCCAA 1024
 QY 781 ATTGACGAGGTTTTTAAACAGAGCTTGGCAGTTGGCCTCAATTTGGAATGGAATTTGCA 840
 DB 1025 ATTGACGAGGTTTTTAAACAGAGCTTGGCAGTTGGCCTCAATTTGGAATGGAATTTGCA 1084
 QY 841 TTGATGAACAGTACATTTGATTTGATTTTATGAGTACAGATTTACTTTGGAACTTTGCA 900
 DB 1085 TTGATGAACAGTACATTTGATTTGATTTTATGAGTACAGATTTACTTTGGAACTTTGCA 1144
 QY 901 AAGGTTTTTTCAGGAGAGAAATCTTTTGTATTTTATGAGTACAGATTTACTTTGGAACTTTGCA 960
 DB 1145 AAGGTTTTTTCAGGAGAGAAATCTTTTGTATTTTATGAGTACAGATTTACTTTGGAACTTTGCA 1204
 QY 961 ACAAAATTTCTTGGAGAAACAGTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAGAAACC 1020
 DB 1205 ACAAAATTTCTTGGAGAAACAGTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAGAAACC 1264
 QY 1021 ACAGATAAGCTTCTTACCTTTGGATGAGATTTT 1053
 DB 1265 ACAGATAAGCTTCTTACCTTTGGATGAGATTTT 1297

RESULT 3

US-10-342-887-71
 ; Sequence 71, Application US/10342887
 ; Publication No. US20040058340A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dai, Hongyue
 ; APPLICANT: He, Yudong
 ; APPLICANT: Linsley, Peter S.
 ; APPLICANT: Mao, Mao
 ; APPLICANT: Roberts, Christopher J.
 ; APPLICANT: Van 't Veer, Laura Johanna
 ; APPLICANT: Van de Vijver, Marc J.
 ; APPLICANT: Bernards, Rene
 ; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
 ; FILE REFERENCE: 9301-188-999
 ; CURRENT APPLICATION NUMBER: US/10/342,887
 ; CURRENT FILING DATE: 2003-01-15
 ; PRIOR APPLICATION NUMBER: 60/298,918
 ; PRIOR FILING DATE: 2001-06-18
 ; PRIOR APPLICATION NUMBER: 60/380,710
 ; PRIOR FILING DATE: 2002-05-14
 ; PRIOR APPLICATION NUMBER: 10/172,118
 ; PRIOR FILING DATE: 2002-06-14
 ; NUMBER OF SEQ ID NOS: 2699
 ; SEQ ID NO 71
 ; LENGTH: 4955
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-342-887-71
 Query Match 100.0%; Score 1053; DB 18; Length 4955;
 Best Local Similarity 100.0%; Pred. No. 2e-294;
 Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGGGCAGCCGGAAGGCGGAGCGGCGGCTGGATCAGAGATCATCTTCA 60
 DB 245 ATGGGGCAGCCGGAAGGCGGAGCGGCGGCTGGATCAGAGATCATCTTCA 304
 QY 61 GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCATCTCTAAGAAAGAGTTCTCGC 120

DB 305 GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCATCTCTAAGAAAGAGTTCTCGC 364
 QY 121 CGGTTTGTCTATCTTTTCAATCAGTACCTCATATTTTGGAAAAATGTATAAACAGGCACAG 180
 DB 365 CGGTTTGTCTATCTTTTCAATCAGTACCTCATATTTTGGAAAAATGTATAAACAGGCACAG 424
 QY 181 GCTTCTCTTGGACAGCAGAGAGGTCGATTTATCAAAAGGATCTCCCTCATCTGGAAACAG 240
 DB 425 GCTTCTCTTGGACAGCAGAGAGGTCGATTTATCAAAAGGATCTCCCTCATCTGGAAACAG 484
 QY 241 CTTAAACAGATGAGAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCGACGCGATGAT 300
 DB 485 CTTAAACAGATGAGAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCGACGCGATGAT 544
 QY 301 GGAATTTGTAATGAAAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTCGAGGTTCCAGAGGCT 360
 DB 545 GGAATTTGTAATGAAAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTCGAGGTTCCAGAGGCT 604
 QY 361 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
 DB 605 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 664
 QY 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAAATGCAATTTGAA 480
 DB 665 CTGATAGACATTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAAATGCAATTTGAA 724
 QY 481 ACCATGCCCCATGTTTAAAGAAAAAGAGATTTGGGCTTTCCGATGATGATGATGATGATGATG 540
 DB 725 ACCATGCCCCATGTTTAAAGAAAAAGAGATTTGGGCTTTCCGATGATGATGATGATGATGATG 784
 QY 541 TCTACTTTTGGGAAAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGATTTTCTTCTCAGAGA 600
 DB 785 TCTACTTTTGGGAAAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGATTTTCTTCTCAGAGA 844
 QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGCTTTATGTCAGGAGATCATTTGTGATGCTGTCAA 780
 DB 845 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGCTTTATGTCAGGAGATCATTTGTGATGCTGTCAA 904
 QY 661 AATGAACCTCATCAGCAGAGATGAAGGCTTCACTGTGACATTTGCTTGCCTGATGTTCCAA 720
 DB 905 AATGAACCTCATCAGCAGAGATGAAGGCTTCACTGTGACATTTGCTTGCCTGATGTTCCAA 964
 QY 721 TACTTAGTAAATAAGGCTTCAAGAGAGAGGCTTCAAGAGAGAGGCTTCAAGAGAGATCATTTGTGATGCTGTCAA 1024
 DB 965 TACTTAGTAAATAAGGCTTCAAGAGAGAGGCTTCAAGAGAGAGGCTTCAAGAGAGATCATTTGTGATGCTGTCAA 1084
 QY 781 ATTGACGAGGTTTTTAAACAGAGAGGCTTGGCAGTTGGCCTCAATTTGGAATGGAATTTGCA 840
 DB 1025 ATTGACGAGGTTTTTAAACAGAGAGGCTTGGCAGTTGGCCTCAATTTGGAATGGAATTTGCA 1084
 QY 841 TTGATGAACAGTACATTTGATTTGATTTTATGAGTACAGATTTACTTTGGAACTTTGCA 900
 DB 1085 TTGATGAACAGTACATTTGATTTGATTTTATGAGTACAGATTTACTTTGGAACTTTGCA 1144
 QY 901 AAGGTTTTTTCAGGAGAGAAATCTTTTGTATTTTATGAGTACAGATTTACTTTGGAACTTTGCA 960
 DB 1145 AAGGTTTTTTCAGGAGAGAAATCTTTTGTATTTTATGAGTACAGATTTACTTTGGAACTTTGCA 1204
 QY 961 ACAAAATTTCTTGGAGAAACAGTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAGAAACC 1020
 DB 1205 ACAAAATTTCTTGGAGAAACAGTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAGAAACC 1264
 QY 1021 ACAGATAAGCTTCTTACCTTTGGATGAGATTTT 1053
 DB 1265 ACAGATAAGCTTCTTACCTTTGGATGAGATTTT 1297

RESULT 4

US-10-698-228-3
 ; Sequence 3, Application US/10698228
 ; Publication No. US20040072253A1
 ; GENERAL INFORMATION:


```
QY 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCATCTTAGCCCTTTTTCAGCCAGTGAT 300
DB 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCATCTTAGCCCTTTTTCAGCCAGTGAT 300
QY 301 GGAATTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGAGTTTCCAGAGGCT 360
DB 301 GGAATTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGAGTTTCCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTTG 420
DB 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTTGAA 480
DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTTGAA 480
QY 481 ACCATGCCCTATCTTAAGAAAAGCAGATTTGGCCCTTGGATGATAGCAGATAGAAA 540
DB 481 ACCATGCCCTATCTTAAGAAAAGCAGATTTGGCCCTTGGATGATAGCAGATAGAAA 540
QY 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGTAGAAAGGAGTTTCTTCTCAGGA 600
DB 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGTAGAAAGGAGTTTCTTCTCAGGA 600
QY 601 TCTTTTCTGCTATATTTCTGGCTTAAAGAAAGAGAGGTTCTTATGCCAGGACTCCTTTTTC 660
DB 601 TCTTTTCTGCTATATTTCTGGCTTAAAGAAAGAGAGGTTCTTATGCCAGGACTCCTTTTTC 660
QY 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACATTTGCTGCTGATAGAGGTTTCTTCCAA 720
DB 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACATTTGCTGCTGATAGAGGTTTCTTCCAA 720
QY 721 TACTTAGTAAATAAGCTTTCAGAAAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA 780
DB 721 TACTTAGTAAATAAGCTTTCAGAAAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA 780
QY 781 ATTGAGCAGAGTTTTTAAACAGAAAGCTTGCAGTTGGCCCTCATTTGGAATCAATTTGCATT 840
DB 781 ATTGAGCAGAGTTTTTAAACAGAAAGCTTGCAGTTGGCCCTCATTTGGAATCAATTTGCATT 840
QY 841 TTGATGAACAGTACATTTGATTTGCTGACAGATTAATTTGGAACATTTGGAATTTGGAATTTCA 900
DB 841 TTGATGAACAGTACATTTGATTTGCTGACAGATTAATTTGGAACATTTGGAATTTGGAATTTCA 900
QY 901 AAGGTTTTTTCAGCAGAAATCTTTTGAATTTTATGGAACATTTTTCAGAGGAAA 960
DB 901 AAGGTTTTTTCAGCAGAAATCTTTTGAATTTTATGGAACATTTTTCAGAGGAAA 960
QY 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAAGAAC 1020
DB 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAAGAAC 1020
QY 1021 ACAGATAACGTTCTTCACTTGGATGAGATTTT 1053
DB 1021 ACAGATAACGTTCTTCACTTGGATGAGATTTT 1053
```

RESULT 6

US-10-698-228-4

```
; Sequence 4, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
```

```
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-4
```

```
Query Match 99.8%; Score 1051.4; DB 18; Length 1081;
Best Local Similarity 99.9%; Pred. No. 2.4e-294;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 ATGGCGCACCGGAAAGCGCGGAAAGCGCGGCTTGAATCAGGATGAGAGATCATCTTCA 60
DB 20 ATGGCGCACCGGAAAGCGGAAAGCGCGGCTTGAATCAGGATGAGAGATCATCTTCA 79
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTTCTTAAGAAAGAGTTCTGC 120
DB 80 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTTCTTAAGAAAGAGTTCTGC 139
QY 121 CGGTTTGTCAATTTTCCAAATCCAGTACCTCATATTTTGAATAATGTAATAACAGGCACAG 180
DB 140 CGGTTTGTCAATTTTCCAAATCCAGTACCTCATATTTTGAATAATGTAATAACAGGCACAG 199
QY 181 GCTTCTCTTGGACAGCAGAGAGGTGACATTTCAAGAGATCTCCCTCAGTGGAAACAG 240
DB 200 GCTTCTCTTGGACAGCAGAGAGGTGACATTTCAAGAGATCTCCCTCAGTGGAAACAG 259
QY 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCAATCTTAGCCCTTTTTCAGCCAGTGAT 300
DB 260 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCAATCTTAGCCCTTTTTCAGCCAGTGAT 319
QY 301 GGAATTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGAGTTTCCAGAGGCT 360
DB 320 GGAATTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGAGTTTCCAGAGGCT 379
QY 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTTG 420
DB 380 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTTG 439
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTTGAA 480
DB 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTTGAA 499
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTTCGATGGATAGCAGATAGAAA 540
DB 500 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTTCGATGGATAGCAGATAGAAA 559
QY 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGCTGTAGAGGATTTTCTTCTCAGGA 600
DB 560 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGCTGTAGAGGATTTTCTTCTCAGGA 619
QY 601 TCTTTTCTGCTATATTTCTGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCCTTTTTC 660
DB 620 TCTTTTCTGCTATATTTCTGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCCTTTTTC 679
QY 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGATGCTTTGCTGCTGATTTCCAA 720
DB 680 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGATGCTTTGCTGCTGATTTCCAA 739
QY 721 TACTTAGTAAATAAGCTTTCAGAAAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA 780
DB 740 TACTTAGTAAATAAGCTTTCAGAAAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA 799
QY 781 ATTGAGCAGAGTTTTTAAACAGAAAGCTTGCAGTTGGCCCTCATTTGGAATCAATTTGCATT 840
DB 800 ATTGAGCAGAGTTTTTAAACAGAAAGCTTGCAGTTGGCCCTCATTTGGAATCAATTTGCATT 859
QY 841 TTGATGAACAGTACATTTGATTTGATGCTGACAGATTTACTTTGTGGAATCTTGGATTTCTCA 900
```

Db 860 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 919
 QY 901 AAGGTTTTTTCAGGACAGAAATCCTTTTGTATTTATGAAACATTTCTTTAGAAAGAAAA 960
 Db 920 AAGGTTTTTTCAGGACAGAAATCCTTTTGTATTTATGAAACATTTCTTTAGAAAGAAAA 979
 QY 961 ACAAATTTCTTGGAAACAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAAGAAC 1020
 Db 980 ACAAATTTCTTGGAAACAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAAGAAC 1039
 QY 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
 Db 1040 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1072

RESULT 7

US-09-925-301-505
 ; Sequence 505, Application US/09925301
 ; Patent No. US20020052308A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 ; FILE REFERENCE: PA106
 ; CURRENT APPLICATION NUMBER: US/09/925,301
 ; CURRENT FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: PCT/US00/05882
 ; PRIOR FILING DATE: 2000-03-08
 ; PRIOR APPLICATION NUMBER: 60/124,270
 ; PRIOR FILING DATE: 1999-03-12
 ; NUMBER OF SEQ ID NOS: 1694
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 505
 ; LENGTH: 1989
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1917)
 ; OTHER INFORMATION: n equals a,t,g, or c
 US-09-925-301-505

Query Match 54.9%; Score 578.4; DB 9; Length 1989;
 Best Local Similarity 75.0%; Pred. No. 9.9e-157;
 Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCACTCTAAGAAAGAGTTCTCGCGGTTTGTCACTTTCCCAATCCAGTACC 148
 Db 280 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTGTCTATCTTCCCGATCGAGTACC 339
 QY 149 CTGATATTTGGAAATGTATAACAGGACAGAGCTTCTCTGACAGCAGAGAGGTGCG 208
 Db 340 ATGATATCTGGCAGATGTATAGAGGACAGAGGCTTCTTTGACCGCGGAGAGGTGG 399
 QY 209 ACTTATCAAGAGTCTCCCTCACTGAAACAAAGCTTAAAGCAGATGAGAGTACTTCATCT 268
 Db 400 ACCTCTCCAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 459
 QY 269 CTCACATCTAGCCTTTTGGAGCCAGTGTGAATGTAATGAAATTTGGTGGAGC 328
 Db 460 CCCATGTTCTGGCTTCTTTTGCAGCAAGCGATGATGATGAAATGAAATCTTGGTGGAGC 519
 QY 329 GCTTTAGTCAGGAGGTGAGGTTCCAGAGGCTCGCTTGTATGCTTCTCAATTTCTCA 388
 Db 520 GATTTAGCCAGAGAGTTCAGATTCAGAGCCGCTGTTTCTATGGCTTCCAAATGGCCA 579
 QY 389 TCAGAGATGTTCATCAGAGATGTACAGTTTGTGTATGATGACACTTACATCAGAGATCCCA 448
 Db 580 TGAAGAACATACATCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 639
 QY 449 AGAAAGGGAATTTTATTTAATGCAATTTGAACCATGCGCTTATGTTAAGAAAAAGCAG 508
 Db 640 AAGAAAGGGAATTTCTCTTCAATGCCATTTGAACCGATGCGCTTGTGTCAAGAGAGCAG 699

QY 509 ATTGGCCCTTGGATGATGATAGAGATAAATCTACTTTTGGGAAAGAGTGTGGCCT 568
 Db 700 ACTGGCCCTTGGCTGATTTGGGACAAAGAGGCTACTATGTTGTAACGCTTTGTAGCCT 759
 QY 569 TTCTCTCTGTAGAAGAGATTTTCTCTCAGGATCTTTTGTCTGTATATTTCTTGGCTAAAGA 628
 Db 760 TTCTCTCAGTGGAGGCAATTTTCTTTCCGGTCTTTTGGCTCGATATTTCTTGGCTCAAGA 819
 QY 629 AGAGAGGTCTTATGCCAGGACTCAGCTTTTCCAAATGAATCATCAGCAGAGATGAAGAC 688
 Db 820 AACGAGGACTGATGCTGGCCTCAGATTTTCTAATGAATTTATTAGCAGAGATGAGGTT 879
 QY 689 TTCACTGTGACTTTGCTTGGCTGATGTTCCAAATCTTAGTAAATAAGCCTTCAGAGAAA 748
 Db 880 TACACTGTGATTTTGTCTGCTGATGTTCAACACCTGGTACACAAACCATCGGAGGAGA 939
 QY 749 GGGTCAGGAGGATCATTTGTTGATGCTCAAAATTTGAGCAGGAGTCTTTTAAACAGAGCCT 808
 Db 940 GAGTAAAGAAATTAATTAATGCTGTTGCGATAGAACAGGAGTTCTCTACTGAGGCT 999
 QY 809 TGCAGTTGGCCTCATTTGGAAATGAATTTGATGAAACAGTACATTCAGTTTGTAG 868
 Db 1000 TGCCTGTGAAGCTCATTTGGGATGAATTTGCACTCTAATGAAGCAATACATTCAGTTTGTGG 1059
 QY 869 CTGACAGATTTACTTTGGAACTTGGATTTCTCAAGGTTTTCAGGCAAGAAATCCTTTTG 928
 Db 1060 CAGACAGACTTATGCTGGAACTGGGTTTATGCAAGGTTTTCAGAGTAGAGAACCATTTG 1119
 QY 929 ATTTTATGAAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
 Db 1120 ACTTTATGGAGATATTTCTGAGGAAAGAACTAACTTTTGGAGAGAGATGAGCG 1179
 QY 989 AGTATCAGCGCTTTTGCAGTTTATGCGAGAAACCCAGATAAGCTTCTCACTTGGATGCG 1048
 Db 1180 AGTATCAGAGGATGCGAGTGTCTCAGTCCACAGAGATTTCTTTTACCTTGGATGCTG 1239
 QY 1049 ATTT 1052
 Db 1240 ACTT 1243

RESULT 8

US-10-084-817-342
 ; Sequence 342, Application US/10084817
 ; Publication No. US20030119009A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Susan Stuart
 ; APPLICANT: Jed G. Nuchtern
 ; APPLICANT: Sharon E. Plon
 ; APPLICANT: Jason M. Shohet
 ; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
 ; FILE REFERENCE: PA-0046 US
 ; CURRENT APPLICATION NUMBER: US/10/084,817
 ; CURRENT FILING DATE: 2002-02-25
 ; PRIOR APPLICATION NUMBER: 60/270,784
 ; PRIOR FILING DATE: 2001-02-23
 ; NUMBER OF SEQ ID NOS: 365
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 342
 ; LENGTH: 2216
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No. US20030119009A1 202239.1
 US-10-084-817-342

Query Match 54.9%; Score 578.4; DB 15; Length 2216;
 Best Local Similarity 75.0%; Pred. No. 9.9e-157;
 Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;
 QY 89 ATGAAGAGCCACTCTAAGAAAGAGTTCTCGCGGTTTGTCTCAATTTCCAGTACC 148

Db 284 AGGATGAGCGCTGCTGAGAGAAAAACCCCGCGCTTTGTGTCATCTTCCCATCGAGTACC 343
 Qy 149 CTGATATTGGAAAATGTATAACAGGCAAGGCTTCTCTGGAAGAGAGAGTGG 208
 Db 344 ATGATATCTGGAGATGTATAAGAGGAGAGGCTTCTTTGGACCGCGGAGGAGTGG 403
 Qy 209 ACTTATCAAGAGATCTCCCTCACTGGAAACAAAGCTTAAAGCAGATGAGAATCTTCACT 268
 Db 404 ACTCTCCAAGGACATTCAGCACTGGGAATCCCTGAAACCGGAGGAGATATTTTATAT 463
 Qy 269 CTCACATCTTAGCGCTTTTTCAGCGCAGTGCATGGAATGTGAATGAAATTTGGTGAGC 328
 Db 464 CCATGTTCTGGCTTTCTTTCAGCAAGCGATGGCATGTAATGAATCTTGGTGAGC 523
 Qy 329 GCTTTAGTCAGAGGTCAGAGTTCCAGAGGCTCGCTGTTTCTATGCTTTCAAAATCTCA 388
 Db 524 GATTTAGCCAAGAGTTCAGATACAGAGCGGCTGTTTCTATGGCTTCAAATGGCCA 583
 Qy 389 TCAGAGATGTTCACTCAGAGATGTACAGTTTGTGTATAGACACTTACATCAGAGATCCCA 448
 Db 584 TGGAAAACATACATCTGAAATGTATAGTCTTCTTATGACACTTACATATAAAGATCCCA 643
 Qy 449 AGAAAAGGGAATTTTATTAATTAATGCAATTCGAACCATGCCCTATGTTAAGAAAAAGCAG 508
 Db 644 AAGAAAGGGAATTTCTCTCAATGCCATTCGAACGATGCCCTGTGTGCAAGAGGCGAG 703
 Qy 509 ATTGGGCTTTCGATGATAGCAGATAGAAAAATCTACTTTGGGGAAGAGTGGTGCCT 568
 Db 704 ACTGGGCTTTCGCTGGAATGGGGAACAAAGAGGCTACCTATGCTGAACGTGTGTAGCCT 763
 Qy 569 TTGCTGCTGTAGAGAGATTTCTTCTCAGGATCTTTTGTGCTATATTTCTGGCTAAAGA 628
 Db 764 TTGCTGCTGAGTGAAGGCAATTTCTTTTCCGGTCTTTTGGCTGATATCTGGCTCAAGA 823
 Qy 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCAAATGNACTCATCAGCAGAGATCAAGGAC 688
 Db 824 AACGAGGACTGATGCTGGGCTCACATTTTCTAATGAACCTTATAGCAGAGATGAGGGTT 883
 Qy 689 TTCACGTGACTTTGCTTGGCTGATGTTCCAACTTGTAAATAAGCCTTTCAGAGAAA 748
 Db 884 TACACTGTGATTTTGTCTGCTGATGTTCAACACTGTTGACAAACCATCGAGGAGA 943
 Qy 749 GGTCAAGGAGATCAATTTGTGATGCTGTCAAAATGAGCAGGAGTTTAAACAGAGCCT 808
 Db 944 GAGTAAGAGAAATAATTAATCAATGCTGTTCCGATAGAACAGGAGTTTCTCACTGAGCCT 1003
 Qy 809 TGGCAGTTGGCCTCATTTGGAATGAAATGCAATTTTGTATGAAACAGTACATTTGTTAG 868
 Db 1004 TGCCTGTGAAGCTCATTTGGGATGAATGCACTCTAATGAAGCAATACATTTGAGTTTGTG 1063
 Qy 869 CTGACAGATTTACTTGTGGAACCTTGGATTTCTCAAGGTTTTCAGGCGAGAAAATCCTTTT 928
 Db 1064 CAGACAGACTTATGCTGGAACCTGGGTTTATGCAAGGTTTTCAGAGTAGAAGCCATTTG 1123
 Qy 929 ATTTATGGAACAATTTCTTTAGAGGAAAAACAAATTTCTTTGAGAAACGAGTTTTCAG 988
 Db 1124 ACTTTATGAGATAATTTCACTGGAAGGAAGACTAATCTTTTGAAGAGAGTAGGCG 1183
 Qy 989 AGTATCAGCGTTTTCAGTATGGCAGAAACCAAGATACGTTCTTCACTTGGATGCGAG 1048
 Db 1184 AGTATCAGAGGATGGGAGTGTCAAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTG 1243
 Qy 1049 ATTT 1052
 Db 1244 ACTT 1247

RESULT 9

US-09-954-456-724
 ; Sequence 724, Application US/09954456
 ; Patent No. US20020115057A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Young, Paul

; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can
 ; TITLE OF INVENTION: Sets
 ; FILE REFERENCE: 689290-76
 ; CURRENT APPLICATION NUMBER: US/09/954,456
 ; CURRENT FILING DATE: 2001-09-18
 ; PRIOR APPLICATION NUMBER: US/60/233,617
 ; PRIOR FILING DATE: 2000-09-18
 ; PRIOR APPLICATION NUMBER: US/60/234,052
 ; PRIOR FILING DATE: 2000-09-20
 ; PRIOR APPLICATION NUMBER: US/60/234,923
 ; PRIOR FILING DATE: 2000-09-25
 ; PRIOR APPLICATION NUMBER: US/60/235,134
 ; PRIOR FILING DATE: 2000-09-25
 ; PRIOR APPLICATION NUMBER: US/60/235,637
 ; PRIOR FILING DATE: 2000-09-26
 ; PRIOR APPLICATION NUMBER: US/60/235,638
 ; PRIOR FILING DATE: 2000-09-26
 ; PRIOR APPLICATION NUMBER: US/60/235,711
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,720
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,840
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,863
 ; PRIOR FILING DATE: 2000-09-27
 ; NUMBER OF SEQ ID NOS: 2276
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 724
 ; LENGTH: 2500
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-954-456-724

Query Match 54.9%; Score 578.4; DB 9; Length 2500;
 Best Local Similarity 75.0%; Pred. No. 1.1e-156;
 Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;
 Qy 89 ATGAAGAGCCACTCTTAAGAAAGAGTTCTCCCGGTTTGTCTATCTTTTCCAATCCAGTACC 148
 Db 397 AGGATGAGCGCTGCTGAGAGAAAACCCCGCGCTTGTCTATCTTCCCATCGAGTACC 456
 Qy 149 CTGATATTTGGAANAATGTATAAACAGGCACAGGTTCTCTTGGACAGCAGAGAGTGG 208
 Db 457 ATGATATCTGGCAGATGTATAAGAGGCGAGAGGCTTCTTTTGGACCGCGAGGAGTGG 516
 Qy 209 ACTTATCAAGAGATCTCCCTCACTGGAAACAAAGCTTAAAGCAGATGAGAATCTTCACT 268
 Db 517 ACCTCTCCAAGGACATTCAGCACTGGGAATCCCTGNAACCCGAGGAGAGATATTTTATAT 576
 Qy 269 CTCACTCTTAGCTTTTTCAGCCAGTGAATGGAATGTAAATGAAAATTTGGTGGAGC 328
 Db 577 CCATGTTCTGGCTTTCTTTCAGCAAGCGATGCGATAGTAAGTAAGAAAATTTGGTGGAGC 636
 Qy 329 GCTTTAGTCAGAGGTCAGGTTCCAGAGGCTCGCTGTTCTTCTATGGCTTTCAAATTTCTCA 388
 Db 637 GATTTAGCCAAGAAGTTTCAGATTACAGAACGCCGCTGTTCTATGGCTTCCAAAATTTGCCA 696
 Qy 389 TCAGAAATGTTCACTCAGAGATGTACAGTTTGTGTATAGACACTTACATCAGAGATCCCA 448
 Db 697 TGGAAAACATACATCTCTGAAATGTATAGTCTTCTTATTTGACACTTACATAAAGATCCCA 756
 Qy 449 AGAAAAGGGAATTTTATTAATTAATGCAATTCGAACCATGCCCTATGTTAAGAAAAAGCAG 508
 Db 757 AAGAAAGGGAATTTCTCTTCAATGCCATTTGAACGATGCCCTTGTGTCAAGAGAAAGGCGAG 816
 Qy 509 ATTGGGCTTTCGATGATAGCAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCT 568
 Db 817 ACTGGGCTTTCGCTGGATTTGGGGAACAAAGGCTACCTATGCTGGAACGTTGTAGGCT 876
 Qy 569 TTGCTGCTGTAGAGGAGTTTCTTCTCAGAGATCTTTTGTGCTATATTTCTGGCTTAAGA 628
 Db 877 TTGCTGCAAGTGAAGGCAATTTCTTTTCCGGTCTTTTTCGCTCGATATTTCTGGCTCAAGA 936

US-09-954-456-1827
; Sequence 1827, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR FILING DATE: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 1827
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1827

Query Match 54.9%; Score 578.4; DB 9; Length 2500;
Best Local Similarity 75.0%; Pred. No. 1.1e-156;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCACTCTTAAGAAAGAGTTCTCGCCGGTTTGTCTCTTCCCAATCCAGPACC 148
DB 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTGTCTCTTCCCAATCCAGPACC 456

QY 149 CTGATATTTGGAATATTAACAGCAGCAGGCTTCTCTGAGCAGCAGAGAGAGTCG 208
DB 457 ATGATATCTGGCAGATGTATAAGAAAGCAGAGGCTTCTCTTGGACCGCGAGAGGTTG 516

QY 209 ACTTATCAAGAGTCTCCCTCACTGGAACAGCTTAAAGCAGATGAGAGTCTTCTATCT 268
DB 517 ACCTCTCAAGAGATCTTCACTGGAATCCCTGAAACCGAGGAGAGATATTTATAT 576

QY 269 CTCACATCTTAGCCTTTTTCAGCCAGTGAATTTGAAATGAAATTTTGGTGAGC 328
DB 577 CCATGTTCTGGCTTCTTTCAGCAGCAGATGCGATGATGAAATCTTGGTGAGC 636

QY 329 GCTTTAGTCAGGAGGTCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATCTCA 388
DB 637 GATTTAGCAAGAAGTTCAGATTACAGAGCCGCTGTTTCTATGGCTTCAAATTTGCCA 696

QY 389 TCGAGATGTTCACTCAGAGATGATACAGTTTGTGTAGATGACACTTACATCAGATCCCA 448
DB 697 TGAAGAACATACATCTTGAATGTATAGTCTTCTTATGACACTTACATGAAGATCCCA 756

QY 449 AGAAAGGGAATTTTATTAATTAAGCAATTTGAACACCTGCTTATGAAAGAAAGCAG 508
DB 757 AGAAAGGGAATTTCTTCAATGCCATGAAACGATGCTTGTGTGCAAGAGAGCAG 816

QY 509 ATTGGGCTTGGATGATGAGATGAGAAATCTACTTTTGGGGAAGAGTGGTGGCCT 568
DB 817 ACTGGGCTTGGCTGGAATGGGGAAGAGAGGCTTACCTATGCTGGAACGTTGTAGCCT 876

QY 569 TTGCTGCTGTAGAAAGAGTTTCTTCTCAGGATCTTTTGTCTCTATATTTCTGCTAAAGA 628
DB 877 TTGCTGCTGTAGAAAGAGTTTCTTCTCAGGATCTTTTGTCTCTATATTTCTGCTCAAGA 936

QY 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGGAC 688
DB 937 AACGAGACTGATGCCCTGGCTCACTTTTCTAATGAACTTATTAGCAGAGATGAGGTT 996

QY 689 TTCACTGTGACTTTTGTCTGCTGATGTTCCAACTATTAGTAATAAAGCCTTCAGAGAAA 748
DB 997 TACACTGTGATTTTGTCTGCTGATGTTCAAACACTGCTGACAAACCATCGAGGAGA 1056

QY 749 GGGTCAAGGAGATCACTTGTGATGCTGTCAAAATTTGAGCAGAGTTTAAACAGAGCCT 808
DB 1057 GAGTAAGAGAAATAATTATCAATGCTGTTCCGATAGAAACAGAGTTCTCTCACTGAGCCT 1116

QY 809 TGCCAGTTGGCTCACTTGGATGATGATTTGATCAATTTGATCAAAACAGTACATTTGATG 868
DB 1117 TGCCTGTGAAGCTCACTTGGGATGATTTGCACTCTTAATGAAGCAATACATTTGATG 1176

QY 869 CTGACAGATTAATTTGTGAACCTTGAATCTCAAGGTTTTCAGGCAGAAAATCCTTTTG 928
DB 1177 CAGACAGCTTATGCTGGAATGCTGTTTGAAGGTTTTCAGAGTAGAGAACCCATTG 1236

QY 929 ATTTTATGGAACAATTTCTTTAGAAGGAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
DB 1237 ACTTTATGGAGATATTTCTCTGGAAGGAAAGACTAACTCTTTTGAAGAGAGTAGGCG 1296

QY 989 AGTATCAGCTTTTGCAGTTATGCGAGAAACCAAGATACGCTTCACTTGGATGCGAG 1048
DB 1297 AGTATCAGAGATGCGAGTGTGTCAAGTCCAACAGAGAAATTTCTTTTACCTTTGATGCTG 1356

QY 1049 ATTT 1052
DB 1357 ACTT 1360

RESULT 12
US-10-641-643-1370
; Sequence 1370, Application US/10641643
; Publication No. US20040077003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; Susan G. Stuart
; Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
; GENE EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/641,643
; FILING DATE: 14-Aug-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:

```

;
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1370:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2500 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g36154
; SEQUENCE DESCRIPTION: SEQ ID NO: 1370 :
US-10-641-643-1370

Query Match      54.9%; Score 578.4; DB 18; Length 2500;
Best Local Similarity 75.0%; Pred. No. 1.1e-156;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCATCTCTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTTCCATCCAGTACC 148
Db 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTCTATCTTTCCCATCGAGTACC 456
QY 149 CTGATATTTGGAAATGTATAACAGGCACAGGCTTCTCTCGACAGAGAGGTCG 208
Db 457 ATGATATCTGGCAGATGTATAAGAGGCGAGAGGCTTCTTTTGGACCGCGAGGAGTTG 516
QY 209 ACTTATCAAAAGGATCTCCCTCACTGGAAACAGCTTAAAGCAGAGATGAGAGTACTCT 268
Db 517 ACCTCTCAAGGACATTCAGCATGCGCATGCGCATGAGTAAATGAAACTTGTGTGAGC 636
QY 329 GCTTTAGTACAGAGGTCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 388
Db 637 GATTTAGCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 696
QY 389 TCGAGAGATGTTCACTCAGAGATGTACAGATTTGTGCTGATGATGATGATGATGATGAT 448
Db 697 TCGAAGACATACATCTCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCA 756
QY 449 AGAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCGCTATGTTTAAAGAAAGAGCAG 508
Db 757 AAGAAAGGGAATTTCTCTCAATGCCATTGAAACGATGCGCTTGTGTCAAGAGAGGCGAG 816

;
; Query Match      54.9%; Score 578.4; DB 20; Length 2500;
; Best Local Similarity 75.0%; Pred. No. 1.1e-156;
; Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;
;
; QY 89 ATCAAGAGCCATCTCTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTTCCATCCAGTACC 148
; Db 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTCTATCTTTCCCATCGAGTACC 456
; QY 149 CTGATATTTGGAAATGTATAACAGGCACAGGCTTCTCTCGACAGAGAGGTCG 208
; Db 457 ATGATATCTGGCAGATGTATAAGAGGCGAGAGGCTTCTTTTGGACCGCGAGGAGTTG 516
; QY 209 ACTTATCAAAAGGATCTCCCTCACTGGAAACAGCTTAAAGCAGAGATGAGAGTACTCT 268
; Db 517 ACCTCTCAAGGACATTCAGCATGCGCATGCGCATGAGTAAATGAAACTTGTGTGAGC 636
; QY 329 GCTTTAGTACAGAGGTCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 388
; Db 637 GATTTAGCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 696
; QY 389 TCGAGAGATGTTCACTCAGAGATGTACAGATTTGTGCTGATGATGATGATGATGATGAT 448
; Db 697 TCGAAGACATACATCTCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCA 756
; QY 449 AGAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCGCTATGTTTAAAGAAAGAGCAG 508
; Db 757 AAGAAAGGGAATTTCTCTCAATGCCATTGAAACGATGCGCTTGTGTCAAGAGAGGCGAG 816

```

```

QY 929 ATTTATGGAACAACATTTCTTTAGAAAGGAAAAACAATTTCTTTGAGAAACGAGTTTCAG 988
Db 1237 ACTTTATGGAATAATTTCTACCTGGAAGGAAAGACTAACTTTCTTTGAGAAGAGAGTAGCG 1296
QY 989 AGTATCAGCGTTTGTGAGTTATGCGAGAAACCAACAGATAACGCTTCCACCTTGGATGCG 1048
Db 1297 AGTATCAGAGGATGGGAGTGATGTCAAGTCCAAACAGAGAATTTCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 13
US-10-733-878-458
; Sequence 458, Application US/10733878
; Publication No. US20040224408A1
; GENERAL INFORMATION:
; APPLICANT: Jean-Philippe Girard
; APPLICANT: Francois Amalric
; APPLICANT: Myriam Roussigne
; APPLICANT: Thomas Clouaire
; TITLE OF INVENTION: THAP PROTEINS AS NUCLEAR RECEPTORS FOR
; TITLE OF INVENTION: CHEMOKINES AND ROLES IN TRANSCRIPTIONAL REGULATION, CELL
; TITLE OF INVENTION: PROLIFERATION AND CELL DIFFERENTIATION
; FILE REFERENCE: BIOBANK.012A
; CURRENT APPLICATION NUMBER: US/10/733,878
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 60/432699
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/485027
; PRIOR FILING DATE: 2003-07-03
; NUMBER OF SEQ ID NOS: 535
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 458
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-733-878-458

```

```

Query Match      54.9%; Score 578.4; DB 20; Length 2500;
Best Local Similarity 75.0%; Pred. No. 1.1e-156;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 89 ATCAAGAGCCATCTCTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTTCCATCCAGTACC 148
Db 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTCTATCTTTCCCATCGAGTACC 456
QY 149 CTGATATTTGGAAATGTATAACAGGCACAGGCTTCTCTCGACAGAGAGGTCG 208
Db 457 ATGATATCTGGCAGATGTATAAGAGGCGAGAGGCTTCTTTTGGACCGCGAGGAGTTG 516
QY 209 ACTTATCAAAAGGATCTCCCTCACTGGAAACAGCTTAAAGCAGAGATGAGAGTACTCT 268
Db 517 ACCTCTCAAGGACATTCAGCATGCGCATGCGCATGAGTAAATGAAACTTGTGTGAGC 576
QY 269 CTCACTCTTAGCGCTTTTTCAGCCAGTGTGAAATGTAATGAAATTTGGTGGAGC 328
Db 577 CCATGTTCTGGCTTTCTTTGACAGCAAGCGATGCAATAGTAAATGAAATTTGGTGGAGC 636
QY 329 GCTTTAGTACAGAGGTCAGGTTCCAGAGGCTCGCTGTTCTATGGCTTTTCAAAATTTCTCA 388
Db 637 GATTTAGCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 696
QY 389 TCGAGAGATGTTCACTCAGAGATGTACAGATTTGTGCTGATGATGATGATGATGATGAT 448
Db 697 TCGAAGACATACATCTCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCA 756
QY 449 AGAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCGCTATGTTTAAAGAAAGAGCAG 508
Db 757 AAGAAAGGGAATTTCTCTCAATGCCATTGAAACGATGCGCTTGTGTCAAGAGAGGCGAG 816

```


Db 1297 AGTATCAGAGGATGGGAGTGATGTCAAGTCCACAGAGAATTTCTTTTACCTTTGGATGCTG 1356
 QY 1049 ATTT 1052
 Db 1357 ACTT 1360

RESULT 15
 US-10-843-641A-4196
 ; Sequence 4196, Application US/10843641A
 ; Publication No. US2005064454A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Avalon Pharmaceuticals, Inc.
 ; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
 ; TITLE OF INVENTION: Signature Gene Sets
 ; FILE REFERENCE: 689290-189
 ; CURRENT APPLICATION NUMBER: US/10/843,641A
 ; CURRENT FILING DATE: 2004-05-12
 ; PRIOR APPLICATION NUMBER: US/09/873,367
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: US/09/954,531
 ; PRIOR FILING DATE: 2001-09-18
 ; PRIOR APPLICATION NUMBER: US/09/954,456
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US/09/962,436
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US/09/962,832
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US/09/964,824
 ; PRIOR FILING DATE: 2001-09-27
 ; PRIOR APPLICATION NUMBER: US/09/967,768
 ; PRIOR FILING DATE: 2001-09-28
 ; PRIOR APPLICATION NUMBER: US/09/968,007
 ; PRIOR FILING DATE: 2001-10-02
 ; PRIOR APPLICATION NUMBER: US/09/969,347
 ; PRIOR FILING DATE: 2001-10-02
 ; PRIOR APPLICATION NUMBER: US/09/969,708
 ; PRIOR FILING DATE: 2001-10-03
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 847
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 4196
 ; LENGTH: 2500
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-843-641A-4196

Query Match 54.9%; Score 578.4; DB 21; Length 2500;
 Best Local Similarity 75.0%; Pred. No. 1.1e-156;
 Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;
 QY 89 ATGAAGAGCCACTCTTAAGAGAGTTCTGCGCGTTTGTCTATCTTCCAAATCCAGTACC 148
 Db 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTGTCTATCTTCCCGCATCCAGTACC 456
 QY 149 CTGATATTTGGAAGATTAACAGGACAGGCTTCTCTGAGCAGCAGAGAGGTCG 208
 Db 457 ATGATATCTGGCAGATGTATAAGAGAGGAGGCTTCTCTTGGACCGCGAGGAGTTG 516
 QY 209 ACTTATCAAGAGATCTCCCTCACTGGAAACAAAGCTTAAAGCAGATGAGAAGTACTTCTATCT 268
 Db 517 ACTCTCCAGGACATTCAGCACTGGGAATCCCTGAACCCGAGGAGAGATATTTATAT 576
 QY 269 CTCACATCTTAGCTTTTTCAGCCAGTGTAGGAATTTGAAATGAAAAATTTGGTGGAGC 328
 Db 577 CCCATGTTCTGCTTTCTTTTTCAGCAAGCGATGGCATAGTAATGAAACTTGGTGGAGC 636
 QY 329 GCTTTAGTCAGAGGTCGAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTCAAAATCTCA 388
 Db 637 GATTTAGCCAAAGATTCAGATTTACAGAGCCCGCTGTTTCTATGGCTTCCAAATGGCCA 696
 QY 389 TCGAGATGTTCACTCAGAGATGTACAGTTTGTGATAGACACTTACATCAGAGATCCCA 448

Db 697 TGAAGAACATACATCTTGAATGTATAGTCTTCTTATGACACTTACATATAAAGATCCCA 756
 QY 449 AGAAAAAGGGAATTTTATTTTAATGCAATTGAAACCATGCCCTATGTTTAAAGAAAAAGCAG 508
 Db 757 AAGAAAGGGAATTTCTCTCAATGCCATTGAAACGATGCCCTTGTGTCAAGAAAGAGGAG 816
 QY 509 ATTGGGCTTTGGATGAGATAGCAGATAAATACTACTTTTGGGGAAGAGTGGTGGCT 568
 Db 817 ACTGGGCTTTGGCTGGATTTGGGACAAAGAGGCTACCTATGTTGTAACGCTGTTGTAGCT 876
 QY 569 TTGCTGCTGTAGAGGAGTTTCTTCTCAGAGATCTTTTGTGCTATATATCTTGGCTTAAGA 628
 Db 877 TTGCTGAGTGAAGGAGCAATTTCTTTCCGGTCTTTTGGCTCGATATCTTGGCTCAAGA 936
 QY 629 AGAGAGCTCTTATGCCAGGACTCACTTTTCCATGAATCACTCATCAGCAGAGATGAAGGAC 688
 Db 937 AAGAGGACTGATGCCCTGGCTCACATTTTCTAATGAATTTATAGCAGAGATGAGGCT 996
 QY 689 TTCACTGTGACTTTGCTTGGCTGATGTTCCTCAATCTTAGTAAATAAGCTTCAAGAGAAA 748
 Db 997 TACACTGTGATTTTGTCTGCTGATGTTCAAAACACCTGGTACACAAACCATCGAGGAGA 1056
 QY 749 GGTTCAGGAGATCATTTGTTGATGCTCTCAAAATTTGAGCAGGAGCTTTTAAACAGAGCT 808
 Db 1057 GAGTAAGAGAAATAATTTATCAATGCTGTTCCGATAGAAACAGGAGTTCTCTCAGGCT 1116
 QY 809 TGCAGTTGGCTCATTTGGAATGAATTCATTTTGTGATGAACACAGTACATTCAGTGTGTAG 868
 Db 1117 TGCCTGTGAAGCTCATTTGGAGTGAATTCGACTCTAATGAGCAATACATTCAGTGTGTG 1176
 QY 869 CTGACAGATTAATCTGTGGAATCTTGGATTTCTCAAAAGTTTTCAGGCAAGAAATCCTTTG 928
 Db 1177 CAGACAGACTTATGCTGGAATCTGGTTTTCAGAGGTTTTCAGAGTAGAGAACCATTTG 1236
 QY 929 ATTTTATGGAACATTTCTTTAGAGGAAAAACAAATTTCTTTTGAAGAACGAGTTTCAG 988
 Db 1237 ACTTTATGGAGAAATATTTCTACTTGAAGGAAAGACTAACTTCTTTTGAAGAGAGTAGGCG 1296
 QY 989 AGTATCAGCGCTTTTGCAGTTTATGSCAGAAACACAGATAAGCTCTTCACTTGGATGAG 1048
 Db 1297 AGTATCAGAGGATGGGAGTGAATGTCAGTCCAAAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTG 1356
 QY 1049 ATTT 1052
 Db 1357 ACTT 1360

Search completed: June 16, 2005, 11:29:37
 Job time : 756 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2005, 06:56:11 ; Search time 4079 Seconds
(without alignments)
9826.349 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1053

Sequence: 1 atggggcagccgaaagcc.....tcaccttgatgcagatttt 1053

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1053	100.0	3292	3	CR617553 full-leng
2	1053	100.0	4748	3	HSMB020233
3	1041	98.9	4650	3	BC042948 Homo sapi
4	1010.6	96.0	3501	3	CR627376 Homo sapi
5	945	89.7	1056	9	AY398973 Homo sapi
6	929.6	88.3	1083	1	AL547501 AL547501
7	829.8	78.8	1008	9	AY398974 Pan trogl
8	793.2	75.3	879	5	BQ441857 AGENCOURT
9	777	73.8	1043	9	AY398975 Mus muscu
10	731.2	69.4	892	7	CN163214 952695 MA
11	728.2	69.2	914	7	CF995079 AGENCOURT
12	710.6	67.5	870	7	CO648380 ILLUMIGEN
13	682.8	64.8	804	6	CD656661 AGENCOURT
14	655.4	62.2	1061	4	BM801298 AGENCOURT
15	654.8	62.2	896	5	BUI70979 AGENCOURT
16	616.4	58.5	685	5	EX951374 DKPZp781E
17	578.4	54.9	1571	3	CR603461 full-leng
18	578.4	54.9	1573	3	CR625489 full-leng
19	578.4	54.9	1582	3	CF608076 full-leng
20	578.4	54.9	1588	3	CR602054 full-leng
21	578.4	54.9	1592	3	CR621427 full-leng
22	578.4	54.9	1600	3	CR604378 full-leng
23	578.4	54.9	1605	3	CR590959 full-leng
24	578.4	54.9	1612	3	CR614990 full-leng

25	578.4	54.9	1613	3	CR609838 full-leng
26	578.4	54.9	1623	3	CR618451 full-leng
27	578.4	54.9	1630	3	CR602150 full-leng
28	578.4	54.9	1796	3	CR596700 full-leng
29	570.4	54.2	2088	3	AK088907 Mus muscu
30	562	53.4	638	7	CN358022 170005831
31	547	51.9	547	7	CN358025 170005326
32	530.8	50.4	836	7	CR416843 CR416843
33	528.8	50.2	608	1	AUI37050 AUI37050
34	527.2	50.1	540	5	BU429508 UI-HF-BNO
35	525	49.9	997	4	BM468712 AGENCOURT
36	518.8	49.3	522	5	BQ694828 1000971 H
37	513.6	48.8	793	7	CK778582 965745 MA
38	503.8	47.8	923	7	CR580780 CR580780
39	503	47.8	792	5	BU328458 603494277
40	502	47.7	510	7	CR540284 DKPZp459N
41	501.4	47.6	697	6	CA328968 UI-M-PVO-
42	493.8	46.9	510	5	EX955728 DKPZp781G
43	493.2	46.8	946	7	CN024619 AGENCOURT
44	491.8	46.7	932	5	BUI96941 AGENCOURT
45	489.4	46.5	909	6	CA981614 AGENCOURT

ALIGNMENTS

RESULT 1
CR617553 3292 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CSODI011YF14 of Placenta Cot 25-normalized
DEFINITION of Homo sapiens (human).
ACCESSION CR617553
VERSION CR617553.1 GI:50498360
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3292)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
REFERENCE 2 (bases 1 to 3292)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES Location/Qualifiers
source
1. .3292
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI011YF14"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

Query Match 100.0%; Score 1053; DB 3; Length 3292;
Best Local Similarity 100.0%; Pred. No. 3.3e-269; Indels 0; Gaps 0;
Matches 1053; Conservative 0; Mismatches 0;

Qy 1 ATGGCGGACCCGGAAGCGCGGCGGCTGGATCAGATGAGATCATCTTCA 60
|||||
Db 27 ATGGCGGACCCGGAAGCGCGGCGGCTGGATCAGATGAGATCATCTTCA 86
|||||

Qy	61	GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC	120
Db	87	GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC	146
Qy	121	CGGTTTGTCTATCTTCCAAATCCAGTACCTCGATATTTGGGAAATGATATAAAGAGGACAG	180
Db	147	CGGTTTGTCTATCTTCCAAATCCAGTACCTCGATATTTGGGAAATGATATAAAGAGGACAG	206
Qy	181	GCCTTCCTTCGGACAGCAGAGAGGTGCGATTCAAGAGATCTCCCTCACTGGAACAAG	240
Db	207	GCCTTCCTTCGGACAGCAGAGAGGTGCGATTCAAGAGATCTCCCTCACTGGAACAAG	266
Qy	241	CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCGAGCCAGTGAT	300
Db	267	CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCGAGCCAGTGAT	326
Qy	301	GGAAATGTAATGAATAATTTGGTGAGCGCTTTAGTCAGAGAGTGCGAGTTCCAGAGCT	360
Db	327	GGAAATGTAATGAATAATTTGGTGAGCGCTTTAGTCAGAGAGTGCGAGTTCCAGAGCT	386
Qy	361	CGCTGTTTCTATGCTTCAAAATCTCATCGAAGTCTTCACTCAGAGATGTACAGTTTG	420
Db	387	CGCTGTTTCTATGCTTCAAAATCTCATCGAAGTCTTCACTCAGAGATGTACAGTTTG	446
Qy	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA	480
Db	447	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA	506
Qy	481	ACCATGCCCTATGTTAAGAAAAAGCAGATGGGCCCTTGCAGTGTAGCAGATAGAAAA	540
Db	507	ACCATGCCCTATGTTAAGAAAAAGCAGATGGGCCCTTGCAGTGTAGCAGATAGAAAA	566
Qy	541	TCTACTTTTGGGAAAGAGTGGTGGCTTCTGCTGTAGAGAGTCTTCTCTCAGGA	600
Db	567	TCTACTTTTGGGAAAGAGTGGTGGCTTCTGCTGTAGAGAGTCTTCTCTCAGGA	626
Qy	601	TCTTTTCTGCTATATCTGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCACCTTTTTC	660
Db	627	TCTTTTCTGCTATATCTGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCACCTTTTTC	686
Qy	661	AATGAATCTATCAGCAGATGAAGACTTCTGCTGACTTTGCTTGCCTGATGTTCCAA	720
Db	687	AATGAATCTATCAGCAGATGAAGACTTCTGCTGACTTTGCTTGCCTGATGTTCCAA	746
Qy	721	TACTTAGTAATAGCCCTTCAGAAAGAGGTGAGGAGATCATTTGATGCTGTCAAA	780
Db	747	TACTTAGTAATAGCCCTTCAGAAAGAGGTGAGGAGATCATTTGATGCTGTCAAA	806
Qy	781	ATTGAGCAGAGTGTATTAACAGAAAGCCTTGCAGTGGCCCTCATTTGGAATGCAATT	840
Db	807	ATTGAGCAGAGTGTATTAACAGAAAGCCTTGCAGTGGCCCTCATTTGGAATGCAATT	866
Qy	841	TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGGAACCTTGGATTCTCA	900
Db	867	TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGGAACCTTGGATTCTCA	926
Qy	901	AAGTTTTCAGGAGAAAAATCCTTTTGAATTTTATGGAACATTTCTTTAGAGGAAAA	960
Db	927	AAGTTTTCAGGAGAAAAATCCTTTTGAATTTTATGGAACATTTCTTTAGAGGAAAA	986
Qy	961	ACAAATTTCTTGGAAACAGTTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC	1020
Db	987	ACAAATTTCTTGGAAACAGTTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC	1046
Qy	1021	ACAGATAACGCTTTCACCTTGGATGACAGATTTT	1053
Db	1047	ACAGATAACGCTTTCACCTTGGATGACAGATTTT	1079

RESULT 2
HSM802033
LOCUS HSM802033 4748 bp mRNA linear HTC 22-SEP-2004

DEFINITION	Homo sapiens mRNA; cDNA DKFZp761E1312 (from clone DKFZp761E1312).
ACCESSION	AL137348
VERSION	AL137348.1 GI:6807859
KEYWORDS	HTC.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	1 (bases 1 to 4748)
AUTHORS	Ansoorge, W., Krieger, S., Regiert, T., Rittmueller, C., Schwager, B., Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.
CONSTRM	The German cDNA Consortium
TITLE	Direct Submission
JOURNAL	Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
COMMENT	Neuberberg, GERMANY Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp761E1312) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp761E1312 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.
FEATURES	Location/Qualifiers
source	1..4748 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="RZPD:DKFZp761E1312" /db_xref="taxon:9606" /clone="DKFZp761E1312" /tissue_type="amygdala" /DH10B; sites NotI + SalI" /dev stage="adult" /note="ribonucleotide reductase"
gene	1..4748 /gene="DKFZp761E1312"
CDS	47..1102 /gene="DKFZp761E1312" /codon_start=1 /product="hypothetical protein" /protein_id="CAB70703.2" /db_xref="GI:52545612" /db_xref="GOA:Q9NTD8" /db_xref="UniProt/TreMBL:Q9NTD8" /translation="MGDPPEAAAGLDQDSSSDTNESEIKSNEEPLLRKSSRRFVI FPQYPIWMYKQASFTAEVDLSKDLPHWNKLKADKVFISHLAPFAADGI VENLVERFSEVQVPEARCFYQGLLENVHSEMYSLIDTVIRDPKKREFLFNAIE TMPTVKKADWLRADIKSTFGFVPAFAVEGSPFSGFAAIFWLKRGMLPGLT FSNELISREGLHCDFACLMFQYLVNKPSEERREIIVDAVKEIEFTEALPVLGIL MNCILMKQYIEFVADRLVLELGFQVFAENPFDPFMEINISLEGKTNFFPKRYSEYQRF AVMAETTDNVFTLDADF"
ORIGIN	
Query Match	100.0%; Score 1053; DB 3; Length 4748;
Best Local Similarity	100.0%; Pred. No. 3.7e-269;
Matches 1053; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 ATGGCGGACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db	47 ATGGCGGACCGCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 106
Qy	61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Db	107 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 166
Qy	121 CGGTTTGTCTATCTTCCAAATCCAGTACCTCGATATTTGGAATGATATAAAGAGGACAG 180
Db	167 CGGTTTGTCTATCTTCCAAATCCAGTACCTCGATATTTGGAATGATATAAAGAGGACAG 226

Qy	181	GCTTCTCTTCGGA	CAGCAGAAAGAGGTGC	ATTATCAAAGGATCTCCCTCACTCGAAACAAG	240
Db	227	GCTTCTCTTCGGA	CAGCAGAAAGAGGTGC	ATTATCAAAGGATCTCCCTCACTCGAAACAAG	286
Qy	241	CTTAAAGCAGATGAGA	GTACTTTCATCTTCACATCTT	TAGCCCTTTTTCGAGCCAGTGAT	300
Db	287	CTTAAAGCAGATGAGA	GTACTTTCATCTTCACATCTT	TAGCCCTTTTTCGAGCCAGTGAT	346
Qy	301	GGAATTCGTAATGAAAA	TTTGGTGGAGCGCTTTAGT	CAGGAGGTGCAGGTTTCCAGAGCCT	360
Db	347	GGAATTCGTAATGAAAA	TTTGGTGGAGCGCTTTAGT	CAGGAGGTGCAGGTTTCCAGAGCCT	406
Qy	361	CGCTGTTTCTATGGCT	TTTCAAATTTCTCATCGAGAAT	TGTTCACTCAGAGATGTGTA	CAGTTTG 420
Db	407	CGCTGTTTCTATGGCT	TTTCAAATTTCTCATCGAGAAT	TGTTCACTCAGAGATGTGTA	CAGTTTG 466
Qy	421	CTGATAGACACTTTA	CATCAGAGATGCCAAGAAA	GGGAAATTTTATTATTAATGCAATTTGAA	480
Db	467	CTGATAGACACTTTA	CATCAGAGATGCCAAGAAA	GGGAAATTTTATTATTAATGCAATTTGAA	526
Qy	481	ACCATGCCCTATGTT	TAAAGAAAAGACAGATT	TGGCCCTTGCAGTAGGATAGCAGATAGAAAA	540
Db	527	ACCATGCCCTATGTT	TAAAGAAAAGACAGATT	TGGCCCTTGCAGTAGGATAGCAGATAGAAAA	586
Qy	541	TCTACTTTTGGGGAAA	GAGTGGTGGCCCTTGTGCT	GTAGAAGAGAGTTTCTTCTCAGAGGA	600
Db	587	TCTACTTTTGGGGAAA	GAGTGGTGGCCCTTGTGCT	GTAGAAGAGAGTTTCTTCTCAGAGGA	646
Qy	601	TCTTTTGTGCTATAT	TCTGGCTAAAGAGAGAGGT	CTTATGCGAGGACTCAGTTTTTTC	660
Db	647	TCTTTTGTGCTATAT	TCTGGCTAAAGAGAGAGGT	CTTATGCGAGGACTCAGTTTTTTC	706
Qy	661	AATGAACCTCATCAG	CAGAGATGAAGGACTTCA	CTGTGACTTTGCTTGCCTGATGTTCCAA	720
Db	707	AATGAACCTCATCAG	CAGAGATGAAGGACTTCA	CTGTGACTTTGCTTGCCTGATGTTCCAA	766
Qy	721	TACTTTAGTAATAGC	CTTCAGAAAGAGGT	CAGGAGATCAATGTTGATGCTGTCAAA	780
Db	767	TACTTTAGTAATAGC	CTTCAGAAAGAGGT	CAGGAGATCAATGTTGATGCTGTCAAA	826
Qy	781	ATTGAGCAGGAGTTT	TTTAAACAGAACGCTT	CGCCAGTTGGCCTCATTTGGAATGAAATTCGATT	840
Db	827	ATTGAGCAGGAGTTT	TTTAAACAGAACGCTT	CGCCAGTTGGCCTCATTTGGAATGAAATTCGATT	886
Qy	841	TTGATGAAA	CAGTACATGTGATTTGT	AGCTGACAGATTTACTTTGTGGAACCTTGGAATTCCTCA	900
Db	887	TTGATGAAA	CAGTACATGTGATTTGT	AGCTGACAGATTTACTTTGTGGAACCTTGGAATTCCTCA	946
Qy	901	AAGGTTTTTTCAGG	CAGAAAATCCCTTTTGATTTT	TATGGAACAAATTTCTTTAGAGGAAAA	960
Db	947	AAGGTTTTTTCAGG	CAGAAAATCCCTTTTGATTTT	TATGGAACAAATTTCTTTAGAGGAAAA	1006
Qy	961	ACAAATTTCTTTG	AGAAACGAGTTT	CAGAGTATCAGCGTTTTG	CAGTATAGGCAGAAACC 1020
Db	1007	ACAAATTTCTTTG	AGAAACGAGTTT	CAGAGTATCAGCGTTTTG	CAGTATAGGCAGAAACC 1066
Qy	1021	ACAGATAACGCT	TTTCACTTTGGATGCA	ATTTTT 1053	
Db	1067	ACAGATAACGCT	TTTCACTTTGGATGCA	ATTTTT 1099	

RESULT 3	BC042948	BC042948	4650 bp	linear	HTC 12-OCT-2004
LOCUS	BC042948				
DEFINITION	BC042948	Homo sapiens cDNA clone IMAGE:4798175, containing frame-shift errors.			
ACCESSION	BC042948				
VERSION	BC042948.2	GI:34194000			
KEYWORDS	HTC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				

REFERENCE
AUTHORS

1 (bases 1 to 4650)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altshul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schetz, T.E., Brownstein, M.J., Ussdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.C., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krywianksi, M.I., Skalska, U., Smalls, D.E., Schnurch, A., Schein, J.E., Jones, S.J. and de Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932

2 (bases 1 to 4650)
Director MGC Project.
Direct Submission
Submitted (09-JAN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

TITLE

JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 25, 2003 this sequence version replaced gi:27695575.
Contact: MGC help desk
Email: cgabps-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran-Guin, Nancy Liao,
Kim MacDonald, Amara Masson, Mike K. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lilisa Prahbu,
Farvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska,
Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie
Schnein, Asim Siddiqui, Rob Holt, Marco Marra.

FRATERNITY

source

```
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/clone IMAGE:4798175"  
/tissue type="Brain, hypothalamus"  
/clone_lib="NIH MGC_96"  
/lab_host="DH10B"  
note=Vector: pBluescript"
```

ORIGIN

```

Query Match      98.9%; Score 1041; DB 3; Length 4650;
Best Local Similarity 99.9%; Pred. No. 5.8e-266;
Matches 1052; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy      1  ATGGGCGACCCGGAAAGCGCCGAAGCGCGCGCTGGATCAGATGAGAGATCATCTTCA 60
      |||||

```


Db 72 ATGGCGACCGGAAAGCGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 131
Qy 61 GACACCAACGAAAGTGAATAAGTCAAAAGTGAAGAGCAGCTCTTCAAGAAAGAGTTCTCGC 120
Db 132 GACACCAACGAAAGTGAATAAGTCAAAAGTGAAGAG-CAGTCTTCAAGAAAGAGTTCTCGC 190
Qy 121 CGTTTGTCTATCTTTTCCAAATCCAGTACCTGATATTTGGAAATGTATAACAGGACAG 180
Db 191 CGTTTGTCTATCTTTTCCAAATCCAGTACCTGATATTTGGAAATGTATAACAGGACAG 250
Qy 181 GCTTCTTCTGAGCAGCAGAGAGGTGAGCTTATCAAAAGAGTCTCCCTCACTGGAACAAG 240
Db 251 GCTTCTTCTGAGCAGCAGAGAGGTGAGCTTATCAAAAGAGTCTCCCTCACTGGAACAAG 310
Qy 241 CTTAAAGCAGATGAGAGTACTTCACTCTCACATCTTAGCTTTTGGAGCCAGTGAT 300
Db 311 CTTAAAGCAGATGAGAGTACTTCACTCTCACATCTTAGCTTTTGGAGCCAGTGAT 370
Qy 301 GGAATTGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTTCCAGAGGCT 360
Db 371 GGAATTGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTTCCAGAGGCT 430
Qy 361 CGCTGTTTCTATGCTTTCAAAATCTCATCGAGATGTTCACTCAGAGATGATCAGTTG 420
Db 431 CGCTGTTTCTATGCTTTCAAAATCTCATCGAGATGTTCACTCAGAGATGATCAGTTG 490
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
Db 491 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 550
Qy 481 ACCATGCCCTATGTTTAAAGAAAAAGCAGATTTGGGCCCTTGCATGATGATGAGATGAGAAA 540
Db 551 ACCATGCCCTATGTTTAAAGAAAAAGCAGATTTGGGCCCTTGCATGATGATGAGATGAGAAA 610
Qy 541 TCTACTTTTGGGAGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGTTTCTTCTCAGGA 600
Db 611 TCTACTTTTGGGAGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGTTTCTTCTCAGGA 670
Qy 601 TCTTTTCTGCTATTTCTGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCACATTTTCC 660
Db 671 TCTTTTCTGCTATTTCTGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCACATTTTCC 730
Qy 661 AATGAACTCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 720
Db 731 AATGAACTCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 790
Qy 721 TACTTAGTAATAAGCCTTTCAGAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA 780
Db 791 TACTTAGTAATAAGCCTTTCAGAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA 850
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGGCAGTGGCCTCATTTGGAATGCAATTGCATT 840
Db 851 ATTGAGCAGGAGTTTAAACAGAGCCTTGGCAGTGGCCTCATTTGGAATGCAATTGCATT 910
Qy 841 TTGATGAAACAGTACATTGATTTGTAGTGCAGAGATTACTTGTGGAACCTTGGATTCTCA 900
Db 911 TTGATGAAACAGTACATTGATTTGTAGTGCAGAGATTACTTGTGGAACCTTGGATTCTCA 970
Qy 901 AAGGTTTTTCAGGAGAGAAATCCTTTTGAATTTTGAAGAAACATTTCTTTAGAGAGGAAA 960
Db 971 AAGGTTTTTCAGGAGAGAAATCCTTTTGAATTTTGAAGAAACATTTCTTTAGAGAGGAAA 1030
Qy 961 ACAAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTTGCAGTTATGGCAGAAC 1020
Db 1031 ACAAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTTGCAGTTATGGCAGAAC 1090
Qy 1021 ACAGATAAGCTTCTTCACTTGGATGCAGATTTT 1053
Db 1091 ACAGATAAGCTTCTTCACTTGGATGCAGATTTT 1123

LOCUS CR627376 3501 bp mRNA linear HTC 22-SEP-2004
DEFINITION Homo sapiens mRNA; cDNA DKFZp686M05248 (from clone DKFZp686M05248).
ACCESSION CR627376
VERSION CR627376.1 GI:50949847
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3501)
AUTHORS Ottenwälder,B., Obermaier,B., Deutschenbaur,S., Schaipp,A.,
Mewes,H.W., Weil,B., Amlid,C., Osanger,A., Pöbo,G., Han,M. and
Wiemann,S.
CONSTRM The German cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Medigenomix (Martinsried/Germany) within the cDNA
sequencing consortium of the German Genome Project.
This clone (DKFZp686M05248) is available at the RZPD Deutsches
Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp686M05248
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.
FEATURES
Location/Qualifiers
1..3501
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686M05248"
/tissue_type="salivary gland"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
/note="ribonucleotide reductase M2 B (TP53 inducible)"
gene 1..3501
/gene="DKFZp686M05248"
332..1225
/gene="DKFZp686M05248"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH10473.1"
/db_xref="GI:50949848"
/translation="MYKQASFTVAEVDLSKDLPHNKLKADKDYFISHILAPFAA
SDGIVNENLVERFQVQVPEARCFYQILNIENVHSEMYSLIDTYIDPKKREPLF
NAITMPYVKKADWALWIAIDRSTFGERVVFAAVEGVSFGSAATFWKKRGLM
PGLTFPSNELISRDEGLHCDLAFQYLVNKPPEERVREIIVDAVKIEQEFTEALPV
GLIGNCILMKQYIEFVADRLLVLFQSKVFAENPFDFMENISLEGKTNPFKRVSE
YQRFVAVNAETDNTVFLDADF"
CDS 1..3501
Query Match 96.0%; Score 1010.6; DB 3; Length 3501;
Best Local Similarity 98.6%; Pred. No. 6.8e-258;
Matches 1019; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Qy 21 GGAAGCGCCGGGCTGGATCAGGATGAGAGATCATCTTCAGACACCAAGAAAGTGAAT 80
Db 190 GGAAGCGCTTGGAGGCTTCAGGCGGGAGATCATCTTCAGACACCAAGAAAGTGAAT 249
Qy 81 AAAGTCAAAATCAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTGTCTATCTTCCAAT 140
Db 250 AAAGTCAAAATCAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTGTCTATCTTCCAAT 309
Qy 141 CCAGTACCCTCATATTTGGAAAATGTATAAACAGGACACAGGCTTCTTCTGGACAGCAGA 200
Db 310 CCAGTACCCTCATATTTGGAAAATGTATAAACAGGACACAGGCTTCTTCTGGACAGCAGA 369
Qy 201 AGAGTCCAGTTATCAAAAGGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTA 260
Db 370 AGAGTCCAGTTATCAAAAGGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTA 429


```
AY398974      AV398974      1008 bp      DNA      linear      GSS 12-DEC-2003
LOCUS      Pan troglodytes HCM0069 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION      genomic survey sequence.
ACCESSION      AV398974
VERSION      AV398974.1 GI:39754963
KEYWORDS      GSS.
SOURCE      Pan troglodytes (chimpanzee)
ORGANISM      Pan troglodytes
REFERENCE      1 (bases 1 to 1008)
AUTHORS      Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
              Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
              Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
              Adams,M.D. and Cargill,M.
TITLE      Inferring nonneutral evolution from human-chimp-mouse orthologous
              gene trios
JOURNAL      Science 302 (5652), 1960-1963 (2003)
PUBMED      14671302
REFERENCE      2 (bases 1 to 1008)
AUTHORS      Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
              Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
              Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
              Adams,M.D. and Cargill,M.
TITLE      Direct Submission
JOURNAL      Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
              Rockville, MD 20850, USA
COMMENT      This sequence was made by sequencing genomic exons and ordering
              them based on alignment.
FEATURES      Location/Qualifiers
              source      1..1008
                      /organism="Pan troglodytes"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9598"
                      <1..>1008
                      /locus_tag="HCM0069"
              gene
              ORIGIN
Query Match      78.8%; Score 829.8; DB 9; Length 1008;
Best Local Similarity      82.7%; Pred. No. 7.6e-210;
Matches      831; Conservative 0; Mismatches 174; Indels 0; Gaps 0;
Qy      49 AGATCATCTTCAGACACCAACGAAAGTGAATAAGTCAATGAAGACCACTCCTAAGA 108
Db      1 AGATCATCTTCAGACACCAACGAAAGTGAATAAGTCAATGAAGACCACTCCTAAGA 60
Qy      109 AAGAGTTCTCGCCGGTTGTTCATCTTTCCATCCAGTACCCTGATATTTGGAAATGTAT 168
Db      61 AAGAGTTCTCGCCGGTTGTTCATCTTTCCATCCAGTACCCTGATATTTGGAAATGTAT 120
Qy      169 AAACAGGACAGGCTTCCTTCGACAGCAGAGAGGTGCACTTATCAAGAGATCTCCCT 228
Db      121 AAACAGGACAGGCTTCCTTCGACAGCAGAGAGNTCGACTTATCAAGAGATCTCCCT 180
Qy      229 CACTGGACAGCTTAAAGCAGATGAGAGTACTTCTCTCATCTCAGCTTACGCTTTT 288
Db      181 CACTNGAACAGCTTAAAGCAGATGAGAGTACTTCTCTCATCTTACGCTTTT 240
Qy      289 GCAGCCAGTGATGAAATTTGAAATGAAATTTGTTGGAGCGCTTTAGTCAGGAGGTGCAG 348
Db      241 GCAGCCAGTGATGAAATTTGAAATGAAATTTGTTGGAGCGCTTTAGTCAGGAGGTGCAG 300
Qy      349 GTTCAGAGGCTCGCTGTTTCATGCGTTTCAAAATCTCATCGAAGATGTTCACTCAGAG 408
Db      301 GTTCAGAGGCTCGCTGTTTCATGCGTTTCAAAATCTCATCGAAGATGTTCACTCAGAG 360
Qy      409 ATGTACAGTTTGTCTGATAGACACTTACATCAGATCCCAAGAAAGGAAATTTTATTT 468
Db      361 ATGTACAGTTTGTCTGATAGACACTTACATCAGATCCCAAGAAAGGAAATTTTATTT 420
Qy      469 AATGCAATTTGAACACCATGCGCTTATGTTTAAAGAAAAAGCAGATTTGGGCTTGGCATGGATA 528
```

```
Db      421 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 480
Qy      529 GCAGATGAAATCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGATT 588
Db      481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540
Qy      589 TTCTTCTCAGGATCTTTTGGTCTGCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGA 648
Db      541 TTCTTCTCAGGATCTTTTGGTCTGCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCNNNN 600
Qy      649 CTCACTTTTCCAAATGAACACTCATCAGCAGATGAAGGACTTCACTGTGACATTTGCTTGC 708
Db      601 NTCATTTTCCAAATGAACACTCANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 660
Qy      709 CTGATGTTCCAAATGAACACTTAAAGCTTTCAGAGAAAGAGTTCAGGAGAGATCATTTGTT 768
Db      661 CTGATGTTCCAAATGAACACTTAAAGCTTTCAGAGAAAGAGTTCAGGAGAGATCATTTGTT 720
Qy      769 GATGCTGTCAAAATTTAGCAGGAGTTTAAACAGAAAGCTTGCAGTTCGCCCTCATTTGA 828
Db      721 GATGCTGTCAAAATTTAGCAGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 780
Qy      829 ATGAATTCATTTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTGTGGA 888
Db      781 ATGAATTCATTTGATGANAAGTACATTTGAGTTTGTAGCTGACAGATTACTTGTGGA 840
Qy      889 CTGATGTTCTCAAAAGTTTTCAGGACAGAAATCTTTGATTTTATGAAAAACATTTCT 948
Db      841 CTGATGTTCTCAAAAGTTTTCAGGACAGAAATNNNTTTGATTTATGAAAAANNNNNNN 900
Qy      949 TTAGAAGGAAAAACAAATTTCTTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTT 1008
Db      901 TTAGAAGGAAAAACAAANNNNNNTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTT 960
Qy      1009 ATGCAGAAACACAGATAAGCTTTCACCTTGGATGAGAGATTT 1053
Db      961 ATGCAGAAACCCAGATAACGNNNTTCCCTTGGATGAGAGATTT 1005

RESULT 8
LOCUS      BQ441857
DEFINITION      AGENCOURT 7827448 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6014179
              5', mRNA sequence.
ACCESSION      BQ441857
VERSION      BQ441857.1 GI:21180933
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 879)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
              National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
JOURNAL
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: rsb@nih.gov
              Tissue Procurement: ATCC
              CDNA Library Preparation: Life Technologies, Inc.
              CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLAM3208 row: d column: 20
              High quality sequence stop: 662.
FEATURES      Location/Qualifiers
              source      1..879
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:6014179"
                      /tissue_type="embryonal carcinoma, cell line"
```

<p>/lab_host="DH10B (phage-resistant)"</p> <p>/clone_lib="NIH_MGC_92"</p> <p>/note="Organ: Testis; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH MGC Library."</p>	<p>DEFINITION Mus musculus HCM0069 gene, VIRTUAL TRANSCRIPT, partial sequence, Genomic survey sequence.</p> <p>ACCESSION AY398975</p> <p>VERSION AY398975.1 GI:39754964</p> <p>KEYWORDS GSS.</p> <p>SOURCE Mus musculus (house mouse)</p> <p>ORGANISM Mus musculus</p>
---	--

ORIGIN

Query Match	75.3%;	Score 793.2;	DB 5;	Length 879;	
Best Local Similarity	98.7%;	Pred. No. 4.2e-200;			
Matches 821;	Conservative	0;	Mismatches 8;	Indels 3;	Gaps 2;
Qy	1	ATCGGGACCGCGAAAGCGCGGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA	60		
Db	22	ATCGGGACCGCGAAAGCGCGGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA	81		
Qy	61	GACACCAAGCAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGATTCTCGC	120		
Db	82	GACACCAAGCAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGATTCTCGC	141		
Qy	121	CGGTTTGTCATCTTTCCAAATCAGTACCTCGATATTGGAATAATGATAAAGCAGCA	180		
Db	142	CGGTTTGTCATCTTTCCAAATCAGTACCTCGATATTGGAATAATGATAAAGCAGCA	201		
Qy	181	GCTTCTCTTGGACAGACAGAGGTGCACTTATCAAAAGATCTCCCTCATGGAACAAG	240		
Db	202	GCTTCTCTTGGACAGACAGAGGTGCACTTATCAAAAGATCTCCCTCATGGAACAAG	261		
Qy	241	CTTAAAGCAGATGAGAAGTACTTCATCTCTCATCTTAGCCCTTTTTCGAGCCAGTGAT	300		
Db	262	CTTAAAGCAGATGAGAAGTACTTCATCTCTCATCTTAGCCCTTTTTCGAGCCAGTGAT	321		
Qy	301	GGAAATTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAGAGGCTT	360		
Db	322	GGAAATTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAGAGGCTT	381		

Query Match	73.8%;	Score 777;	DB 9;	Length 1043;
Best Local Similarity	81.2%;	Pred. No. 9,2e-196;		
Matches 854; Conservative	0;	Mismatches 185; Indels	13;	Gaps
Qy	1	ATGGGCGACCCGGAAAGGCCGGAGCGCGCGGTGGATCAGATGAGAGATCATCTTCA	60	
Dd	1			
Qy	61	GACACCACGAAGTGAAATAAAGTCAAATGAAGAGCCACTCTTAAGAAGAAGTTCTTCG	120	
Dd	61			
Qy	121	CGGTTTGTCACTCTTTCCAATCCAGTACCTGATATTGGAAAATGTATAAACAGGCACAG	180	
Dd	121			
Qy	181	GCTTCTTCTGACAGCAGAGAAGGTGCACTTATCAAAGGATCTCCCTCACCTGGACAAG	240	
Dd	181			
Qy	241	CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCCCTTTTTTCGAGCCAGTGAT	300	
Dd	241			
Qy	301	GGAAATTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT	360	
Dd	301			
Qy	361	CGCTGTTTCTATGGCTTTCAAATTCTATCGAGAATGTTTCACTCAGAGAGTGACAGTTTG	420	
Dd	361			
Qy	421	CTGATAGACACTTATACAGAGATGCCAAGAAAAAGGGAATTTTATTTTAAATGCCAATTGAA	480	
Dd	421			
Qy	480	CTAAATAGACACTTATACAGAGATCCAAAGAAAAAGGNNNNNNNNNNNNNNNNNNNNN	480	
Dd	480			

RESULT 9
AY398975
LOCUS

Accession	Length	Library	Source
AY398975	1043 bp	DNA	linear
GSS 12-DEC-2003			

DEFINITION AGENCOURT_15621478 NIH MGC_147 Homo sapiens cDNA clone
IMAGE:30528806 5', mRNA sequence.

ACCESSION CF995079
VERSION CF995079.1 GI:38511139
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 914)
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabs-x@mail.nih.gov
Tissue Procurement: Dr. Stefan Haneson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NDAM615 row: f column: 15
High quality sequence stop: 677.

FEATURES
source Location/Qualifiers
1..914
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30528806"
/tissue_type="Human Placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH MGC 147"
/notes="Organ: placenta; Vector: pBluescriptR; Site 1:
ali-XhoI; Site 2: BamHI; Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: This is
a NIH_MGC library."

ORIGIN
Query Match 69.2%; Score 728.2; DB 7; Length 914;
Best Local Similarity 99.1%; Pred. No. 9.1e-183;
Matches 764; Conservative 0; Mismatches 3; Indels 4; Gaps 3;
QY 1 ATGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
DB 87 ATGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA 146
QY 61 GACACCAACCAAGTGAATAAGTCAATTAAGAGCCACTCTTACAGAGTTCCTCGC 120
DB 147 GACACCAACCAAGTGAATAAGTCAATTAAGAGCCACTCTTACAGAGTTCCTCGC 206
QY 121 CGGTTTGTCTATCTTTCCAAATCCAGTACCTCTGATATTTGGAAATGTATTAACAGGCACAG 180
DB 207 CGGTTTGTCTATCTTTCCAAATCCAGTACCTCTGATATTTGGAAATGTATTAACAGGCACAG 266
QY 181 GCTTCTCTTGGACAGCAGAGAGGTGCTATTATCAAAAGATCTCCCTCTCACTGGAAACAAG 240
DB 267 GCTTCTCTTGGACAGCAGAGAGGTGCTATTATCAAAAGATCTCCCTCTCACTGGAAACAAG 326
QY 241 CTTAAGCAGATGAGAAGTACTTCTCTCATCTTACGCTTTTTCAGCCAGCATGAT 300
DB 327 CTTAAGCAGATGAGAAGTACTTCTCTCATCTTACGCTTTTTCAGCCAGTGTAT 386
QY 301 GGAATTGTAATGAAATTTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360

Db 387 GGAATTGTAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGT 446
QY 361 CGCTGTTTCTATGCGCTTTCAAATTTCTCATCAGAGATGTTTCACTCAGAGATGTACAGTTTG 420
Db 447 CGCTGTTTCTATGCGCTTTCAAATTTCTCATCAGAGATGTTTCACTCAGAGATGTACAGTTTG 506
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTAAATGCAATTGAA 480
Db 507 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTAAATGCAATTGAA 566
QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCTTTCGATGATGATGAGATGAGAAAA 540
Db 567 ACCATGCCCTATGTTAAG--AAAAAGCAGATTGGGCTTTCGATGATGATGAGATGAGAAAA 624
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTCCTGCTGTAGAGAGGATTTTCTTCTCAGGA 600
Db 625 TCTACTTTTGGGAAAGAGTGGTGGCTTTCCTGCTGTAGAGAGGATTTTCTTCTCAGGA 684
QY 601 TCTTTTCTGCTATATTCTTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACATTTTTC 660
Db 685 TCTTTTCTGCTATATTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACATTTTTC 744
QY 661 AATGAATCTATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGCTGATGTTCAA 720
Db 745 AATGAATCTATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGCTGATGTTCAA 804
QY 721 TACTTACTTAATAGCC--TTCAGAGAAAGGCTC-AGGAGATCATCTTCTTG 769
Db 805 TACTTACTTAATAGCCCTTTTCAAAAGAAAGGCTCTAGGAAGATCATCTTCTTG 855

RESULT 12

CO648380

LOCUS

DEFINITION

ILLUMIGEN MCQ_41729 Katze_MBP2 Macaca mulatta cDNA clone

IBIUM:25149 5' similar to Bases 5 to 770 highly similar to human

RM2B (Hs.512592), mRNA sequence.

CO648380

CO648380.1 GI:50569874

EST.

Macaca mulatta (rhesus monkey)

Macaca mulatta

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;

Cercopitheidae; Macaca.

1 (bases 1 to 870)

Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magness,C.L.

Large-scale Rhesus Macaque cDNA Sequencing

Unpublished (2003)

Contact: C. Magness

Illumigen Biosciences Inc.

2203 Airport Way S, Suite 450, Seattle, WA 98134, USA

Tel: 2063780400

Fax: 2063780408

Email: cmagness@illumigen.com

Sequenced on 2004.06.02 691 020 bases. Library Preparation: Prof.

Michael Katze lab at University of Washington DNA Sequencing:

Illumigen Biosciences Inc. For further information, see

<http://www.macaque.org>

PCR Primers

FORWARD: CCCTCACTAAAGGGAACAAA

BACKWARD: CACTATAGGGCGAATTGGTGA

Insert length: 870 Std Error: 0.00

Plate: C1000348 row: B column: 06

Seq primer: CCCTCACTAAAGGGAACAAA

POLYA=No.

Location/Qualifiers

1..870

/organism="Macaca mulatta"

/mol_type="mRNA"

/strain="Indian"

/db_xref="taxon:9544"

FEATURES

source

/clone="IBIUN:25149"
/sex="male"
/tissue_type="blood"
/cell_type="PBMc"
/dev_stage="adult"
/lab_host="Electromax DH10B"
/clone_lib="Katz MPB2"
/notes="vector: pDONR 222; Site 1: BsrG I;
Created from CloneMiner cDNA Library Construction kit
(catalog #18249-029)"

ORIGIN

Query Match 67.5%; Score 710.6; DB 7; Length 870;
Best Local Similarity 98.8%; Pred. No. 4.4e-178;
Matches 716; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 329 GCTTTAGTCAGGAGTCCAGGTTCCAGAGGTCGCTGTTTCTATGGCTTCAATTCTCA 388
DB |||||
3 GGTTTAGTCAGGAGTCCAGGTTCCAGAGGTCGCTGTTTCTATGGCTTCAATTCTCA 62
QY 389 TCAGAGATGTTCACTCAGAGATGTACAGTTTGTGTAGACACTTACATCAGATCCCA 448
DB |||||
63 TCAGAGATGTTCACTCAGAGATGTACAGTTTGTGTAGACACTTACATCAGATCCCA 122
QY 449 AGAAAGGGAAATTTTATTAATGCAATGAAACCATGCCCTATGTTAAGAAAAAGCAG 508
DB |||||
123 AGAAAGGGAAATTTTATTAATGCAATGAAACCATGCCCTATGTTAAGAAAAAGCAG 182
QY 509 ATTGGGCTTCGATGATGATAGATAGAAATCTATTTGGGAAAGATGGTGCCT 568
DB |||||
183 ATTGGGCTTCGATGATGATAGATAGAAATCTATTTGGGAAAGATGGTGCCT 242
QY 569 TTGCTGCTGAGAGGATTTTCTCTCAGATCTTTTGTGTATATTTGCTCAAGA 628
DB |||||
243 TTGCTGCTGAGAGGATTTTCTCTCAGATCTTTTGTGTATATTTGCTCAAGA 302
QY 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCATGAACATCATCAGCAGAGATGAAGAC 688
DB |||||
303 AGAGAGTCTTATGCCAGGACTCACTTTTCCATGAACATCATCAGCAGAGATGAAGAC 362
QY 689 TTCACTGATCTTGTGCTGATGTTTCAATCTAGTAAATAAGCCTTCAGAGAA 748
DB |||||
363 TTCACTGATCTTGTGCTGATGTTTCAATCTAGTAAATAAGCCTTCAGAGAA 422
QY 749 GGTTCAGGAGATCATTTGTGCTCTCAAAATGAGCAGGATTTTCAAGAGCCT 808
DB |||||
423 GGTTCAGGAGATCATTTGTGCTCTCAAAATGAGCAGGATTTTCAAGAGCCT 482
QY 809 TGCAGTTGGCCTCATTTGAATGAATGCTATTTTGAATGAACAGTACATTTAGTTGTAG 868
DB |||||
483 TGCAGTTGGCCTCATTTGAATGAATGCTATTTTGAATGAACAGTACATTTAGTTGTAG 542
QY 869 CTCACAGATTACTTGTGGAATCTTGGATCTCAAGGTTTTCAGGCGAGAAATCCTTTTG 928
DB |||||
543 CTCACAGATTACTTGTGGAATCTTGGATCTCAAGGTTTTCAGGCGAGAAATCCTTTTG 602
QY 929 ATTTTGAAGAAATCTTTTGAAGAGAAACAAATTTCTTTGAGAAACGATTTTCAG 988
DB |||||
603 ATTTTGAAGAAATCTTTTGAAGAGAAACAAATTTCTTTGAGAAACGATTTTCAG 662
QY 989 AGTATCAGCGTTTGTGAGTTATGCGAGAAACCAAGATACGCTTCCACCTTGGATGAG 1048
DB |||||
663 AGTATCAGCGTTTGTGAGTTATGCGAGAAACCAAGATACGCTTCCACCTTGGATGAG 722
QY 1049 ATTTT 1053
DB |||||
723 ATTTT 727

RESULT 13

CD656661
LOCUS
DEFINITION AGENCOURT_14555875 NIA Human H1 Embryonic Stem Cell cDNA Library

CD656661 804 bp mRNA linear EST 18-JUN-2003
AGENCOURT_14555875 NIA Human H1 Embryonic Stem Cell cDNA Library

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

(Long) Homo sapiens cDNA clone IMAGE:30427288 5', mRNA sequence.
CD656661
EST.
GI:31897180

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.

1 (Bases 1 to 804)

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cagpbp@mail.nih.gov

Tissue Procurement: Irene Ginis and Mahendra Rao, NIA

cDNA Library Preparation: Yulan Piao and Minoru Ko

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC c lone distribution information

can be found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov

Plate: NDAMS14 row: h column: 17

High quality sequence stop: 680.

FEATURES

source

Location/Qualifiers
1..804
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30427288"
/tissue_type="Embryonic Stem cells"
/cell_line="WA01"
/lab_host="DHI0B (T1 phage-resistant)"
/clone_lib="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)"

/note="Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI;
This is a long-transcript enriched cDNA library (Genome
Res. 11: 1553-1558 (2001). [PMID: 1154199]) from WA01
cell line. Undifferentiated human ES cell line WA01/H1
was obtained from WiCell Research Institute, Inc.,
Madison, WI, cultured according to their instructions, on
MEF feeders. They formed round colonies with defined edges
and were positive for alkaline phosphatase, SSEA-4, OCT3,
OCT4, REX1, UTR, TERT, SOX2, CX43 and CX45. They are
negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLK3, SSEA-1,
TUBB3, NES, GFAP, and EOMES. When confluent (18-10 days
after plating), the ES cells from 4 X 6cm dishes were
treated with 1 mg/ml collagenase, type IV
(Invitrogen/GIBCO) for 5-10 min and gently scraped off
with 5 ml pipette. RNA was purified with Trizol Reagent
from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558
(2001). [PMID:1154199]) Double-stranded cDNAs were
synthesized with an Oligo(dT) primer [Invitrogen:
5'-pGACTAGTCTTAGAGCGAGCGCGCCCTTTTCTTTT-3'] from
3.4g of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to lone-linker IL-Sal4, purified by phenol/chloroform
extraction, and separated from free linkers by
Centricon-100 column. Then, the cDNAs were amplified by
long-range high fidelity PCR using Ex Taq polymerase
(Takara) with a primer Sal4-S for 25 cycles. The products
were purified by phenol/chloroform extraction and
Centricon-100 column. The cDNAs were digested with SalI
and NotI enzymes and cloned into SalI/NotI site of
pCMV-Sports6 plasmid vector. The average insert size is
about 3.6kb."

ORIGIN

Query Match 64.8%; Score 682.8; DB 6; Length 804;
Best Local Similarity 97.3%; Pred. No. 1.1e-170;
Matches 716; Conservative 0; Mismatches 17; Indels 3; Gaps 2;


```

Db      744  CATGAGCTCATCCACGAGATGAGGAGCTTCCCTGTGAACTTTCCTTGCCTTGCTGATGGT 803
Qy      716  TCCAATACTTAGTAA 731
Db      804  TCCAATACTWAAGAA 819

RESULT 15
LOCUS   BUI70979
DEFINITION BUI70979 896 bp mRNA linear EST 04-SEP-2002
          AGENCOURT 7940338 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6143754
          5', mRNA sequence.
ACCESSION BUI70979
VERSION   BUI70979.1 GI:22684963
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 896)
AUTHORS  NIH-MGC http://mgc.nci.nih.gov/.
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished (1999)
COMMENT  Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: ATCC
          cDNA Library Preparation: Life Technologies, Inc.
          DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
          Cloning: Agencourt Bioscience Corporation
          Cloned through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LIA013467 Row: k Column: 19
          High quality sequence stop: 619.
FEATURES             Location/Qualifiers
     source           1..896
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6143754"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_67"
                     /notes="Organ: eye; Vector: pCMV-SPORT6; Site 1: Not I;
                     Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.75 kb. Library constructed by Life
                     Technologies."
ORIGIN
Query Match      62.2%; Score 654.8; DB 5; Length 896;
Best Local Similarity 98.7%; Pred. No. 3.4e-163;
Matches 671; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

Qy      1  ATGGGGACCCGGAAAGCGCGGAGCGCGGGCTGGATCAGATGAGATCATCTTCA 60
Db      55  ATGGGGACCCGGAAAGCGCGGAGCGCGGGCTGGATCAGATGAGATCATCTTCA 114
Qy      61  GACACCAACGAAAGTGAATAAATGAAGTGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db      115  GACACCAACGAAAGTGAATAAATGAAGTGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 174
Qy      121  CGGTTTGTGATCTTTCCATCCAGTACCTGATATTGGAAATGTATTAACAGGCACAG 180
Db      175  CGGTTTGTGATCTTTCCATCCAGTACCTGATATTGGAAATGTATTAACAGGCACAG 234
Qy      181  GCTTCTCTTCGGACAGCAGAGAGGTGCACTTATCAAAAGGATCTCCCTCACTGGAACAAG 240
Db      235  GCTTCTCTTCGGACAGCAGAGAGGTGCACTTATCAAAAGGATCTCCCTCACTGGAACAAG 294
Qy      241  CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCTTTTTCGAGCCAGTGAT 300
Db      295  CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCTTTTTCGAGCCAGTGAT 354

```

```

Qy      301  GGAATGTGAATGAAAAATTTGGTGAGCGCTTTAGTCAAGAGGTGAGGTTCCAGAGGCT 360
Db      355  GGAATGTGAATGAAAAATTTGGTGAGCGCTTTAGTCAAGAGGTGAGGTTCCAGAGGCT 414
Qy      361  CGCTGTTTCTATGGCTTTCAAATTCCTCATCCGAGATGTTCACTCAGAGATGTACAGTTTG 420
Db      415  CGCTGTTTCTATGGCTTTCAAATTCCTCATCCGAGATGTTCACTCAGAGATGTACAGTTTG 474
Qy      421  CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTTAATGCAATTGAA 480
Db      475  CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTTAATGCAATTGAA 534
Qy      481  ACCATGCCCTATGTTAAGAAAAAGCAGATGGGCTTTCGATGAGATGAGATAGAAAA 540
Db      535  ACCATGCCCTATGTTAAGAAAAAGCAGATGGGCTTTCGATGAGATGAGATAGAAAA 594
Qy      541  TCTACTTTTGGGAAAGAGTGGTGCCCTTTCGCTGTAGAGAGGTTTCTTCTCAGGA 600
Db      595  TCTACTTTTGGGAAAGAGTGGTGCCCTTTCGCTGTAGAGAGGTTTCTTCTCAGGA 654
Qy      601  TCTTTTGTGCTATATTCTGGCTAAAGAGAGAGGTCTTTATGCCAGGACTCACTTTTTC 658
Db      655  TCTTTTGTGCTATATTCTGGCTAAAGAGAGAGGTCTTTATGCCAGGACTCACTTTTTC 714
Qy      659  CCAATGAACTCATCAGCAGA 678
Db      715  CAATGAACTCATCAGCAGA 734

```

Search completed: June 16, 2005, 09:39:41
Job time : 4085 secs


```
QY 212 TATCAAGAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTCACTCTC 271
DB 1 TATCAAGAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTCACTCTC 60
QY 272 ACATCTTAGC-CTTTTTCAGCCAGTGTGGAATGTAATGAAATTTGGTGGAGCGC 330
DB 61 ACATCTTAGCCTTTTTCAGCCAGTGTGGAATGTAATGAAATTTGGTGGAGCGC 120
QY 331 TTTAGTCAGGAGTGCAGGTTTCAGAGGCTCGCTTCTATGGCTTTCAAAATCTTCATC 390
DB 121 TTTAGTCAGGAGTGCAGGTTTCAGAGGCTCGCTTCTATGGCTTTCAAAATCTTCATC 180
QY 391 GAGAATGTTTCACTCAGAGATGTACAGTTTCTGATAGACACTTACATCAGAGATCCCAAG 450
DB 181 GAGAATGTTTCACTCAGAGATGTACAGTTTCTGATAGACACTTACATCAGAGATCCCAAG 240
QY 451 AAAAGGGAATTTTATTAATGCAATTTGAAACCAATGCCCTATGTTAAGAAAAAGCAGAT 510
DB 241 AAAAGGGAATTTTATTAATGCAATTTGAAACCAATGCCCTATGTTAAGAAAAAGCAGAT 300
QY 511 TGGGCTTGGGATGATAGAGATAGAAATCTACTTTTGGGGAAGAGTGGGCTTT 570
DB 301 TGGGCTTGGGATGATAGAGATAGAAATCTACTTTTGGGGAAGAGTGGGCTTT 360
QY 571 GCTGCTGTAGAGAGTTTCTTCTCAGAGATCTTTTCTGCTATATTTCTGGCTAAAGAG 630
DB 361 GCTGCTGTAGAGAGTTTCTTCTCAGAGATCTTTTCTGCTATATTTCTGGCTAAAGAG 420
QY 631 AGAGTCTTATGTCAGAGTCACTTTTTCATGAACTCATCAGCAGAGATGAAGACTT 690
DB 421 AGAGTCTTATGTCAGAGTCACTTTTTCATGAACTCATCAGCAGAGATGAAGACTT 480
QY 691 CATGTGACTTTGCTGCTGATGCTTCCCAATCTTACTTAAAGCTTCAAGAGG 750
DB 481 CATGTGACTTTGCTGCTGATGCTTCCCAATCTTACTTAAAGCTTCAAGAGG 540
QY 751 GTCAGGAGATCAATGTTGATGCTGTCAAAATTCAGCAGAGTTTTAAACAGAGCCTTG 810
DB 541 GTCAGGAGATCAATGTTGATGCTGTCAAAATTCAGCAGAGTTTTAAACAGAGCCTTG 600
QY 811 CCAGTTGGCTCAATGGAATGAATGTCAATTTGATGAAAAGTACATGATGTTTGTAGCT 870
DB 601 CCAGTTGGCTCAATGGAATGAATGTCAATTTGATGAAAAGTACATGATGTTTGTAGCT 660
QY 871 GACAGATTAATCTGGAACTTGAATTTCTCAAGGTTTTCAGCAGAGAAATCCTTTTGTAT 930
DB 661 GACAGATTAATCTGGAACTTGAATTTCTCAAGGTTTTCAGCAGAGAAATCCTTTTGTAT 718
QY 931 TTTATGGAACAATTT 946
DB 719 TATGGAACAATTTT 734
```

```
RESULT 14
BM801298 1061 bp mRNA linear EST 05-MAR-2002
LOCUS AGENCOURT_6422860 NIH_MGC_92 Homo sapiens cdna clone IMAGE:5586693
DEFINITION 5', mRNA sequence.
ACCESSION BM801298
VERSION BM801298.1 GI:19118121
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1061)
NIH-MGC http://mgi.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-1@mail.nih.gov
Tissue Procurement: ATCC
```

cdna Library Preparation: Life Technologies, Inc.
cdna Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
<http://image.llnl.gov>
Plate: LHAM12354 row: p column: 22
High quality sequence stop: 631.
Location/Qualifiers

FEATURES

source

1. 1061

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5586693"

/tissue type="embryonal carcinoma, cell line"

/lab host="DH10B (phage-resistant)"

/clone lib="NIH_MGC_92"

/notes="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 2.5 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

ORIGIN

```
Query Match 62.2%; Score 655.4; DB 4; Length 1061;
Best Local Similarity 97.6%; Pred. No. 2.4e-163;
Matches 718; Conservative 0; Mismatches 12; Indels 6; Gaps 5;

QY 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTCA 60
DB 85 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTCA 144
QY 61 GACACCAACCAAGTGAATTAAGTCAAAATCAAGAGCCACTCTTAAGAAAAAGTTCGCG 120
DB 145 GACACCAACCAAGTGAATTAAGTCAAAATCAAGAGCCACTCTTAAGAAAAAGTTCGCG 204
QY 121 CGGTTTCTCATCTTTTCAATCCAGTACCTGATATTTGAAAAATGTATAACAGGACAG 180
DB 205 CGGTTTCTCATCTTTTCAATCCAGTACCTGATATTTGAAAAATGTATAACAGGACAG 264
QY 181 GCTTCTCTTGGACAGCAGAGAGTGCATTTATCAAGGATCTCCCTCAGTGAACAG 240
DB 265 GCTTCTCTTGGACAGCAGAGAGTGCATTTATCAAGGATCTCCCTCAGTGAACAG 324
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCAATCTTAGCCTTTTTCAGCCAGTGTAT 300
DB 325 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCAATCTTAGCCTTTTTCAGCCAGTGTAT 384
QY 301 GGAATTCGTAATGAAAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAGGTTCCAGAGCT 360
DB 385 GGAATTCGTAATGAAAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAGGTTCCAGAGCT 444
QY 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTCTCATCAGAGATGTACAGTTTG 420
DB 445 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTCTCATCAGAGATGTACAGTTTG 504
QY 421 CTGATAGACACTTACATCAGAGATGCCAAGAAAAAGGGAATTTTATTTAAATGCAATTCGA 480
DB 505 CTGATAGACACTTACATCAGAGATGCCAAGAAAAAGGGAATTTTATTTAAATGCAATTCGA 564
QY 481 ACCATGCCCTATGTTTAAAGAAAAAGCAGATTTGGGCTTTGCCGATGGATAGCAGATAGAAA 540
DB 565 ACCATGCCCTATGTTTAAAGAAAAAGCAGATTTGGGCTTTGCCGATGGATAGCAGATAGAAA 624
QY 541 TCTACTTTTGGGAAAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGTTTCTTCTCAGGA 600
DB 625 TCTACTTTTGGGAAAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGTTTCTTCTCAGGA 684
QY 601 TCTTTTCTGCTATATTTCTGGCTAAAGAGAGGCTTTATGCCAGGACTCACTTTTTC 660
DB 685 TCTTTTCTGCTATATTTCTGGCTAAAGAGAGAGGCTTATGCCAGGACTCACTTTTTC 743
QY 661 AATGAACCTCAT-CAGCAGAGAT-GAAGGACTTCACTGTGA--CTTTGCTTGGCTCAT-GT 715
```

```
Qy 261 CTTCACTCTCACATCTTAGCTTTTTCAGCAGATGAGAAATGTAATGAAAATTT 320
Db |||
Qy 430 CTTCACTCTCACATCTTAGCTTTTTCAGCAGATGAGAAATGTAATGAAAATTT 489
Db |||
Qy 321 GGTGAGCGCTTTAGTCAGAGGTCAGGTTCCAGAGCGCTCGCTTTCTATGCTTTCA 380
Db |||
Qy 490 GGTGGAACGCTTTAGTCAGAGGTCAGGTTCCAGAGCGCTCGCTTTCTATGCTTTCA 549
Db |||
Qy 381 AATTCTCATCAGAAATGTTCACTCAGAGATGACAGATTTGCTGTATAGACACTTACATCAG 440
Db |||
Qy 441 AGATCCCAAGAAAAGGAAATTTTATTAATGCAATTGAAACCAATGCCCTATGTAAGAA 500
Db |||
Qy 610 AGATCCCAAGAAAAGGAAATTTTATTAATGCAATTGAAACCAATGCCCTATGTAAGAA 669
Qy 501 AAAAGCAGATGGCGCTTGCGATGATAGCAGATAGAAAATCTACTTTTGGGGAAGAGT 560
Db |||
Qy 670 AAAAGCAGATGGCGCTTGCGATGATAGCAGATAGAAAATCTACTTTTGGGGAAGAGT 729
Qy 561 GGTGGCCTTTGCTGCTGAGAGGAGTTTCTCTCAGAGATCTTTTGTGCTATATCTG 620
Db |||
Qy 730 GGTGGCCTTTGCTGCTGAGAGGAGTTTCTCTCAGAGATCTTTTGTGCTATATCTG 789
Qy 621 GCTAAAGAGAGAGGCTTTATGCCAGGACTCACCTTTTCCAAATGAACTCATCAGCAGAGA 680
Db |||
Qy 790 GCTAAAGAGAGAGGCTTTATGCCAGGACTCACCTTTTCCAAATGAACTCATCAGCAGAGA 849
Qy 681 TGAAGGACTTCACGTGACCTTTGCTGCTGAGATTTCAATACCTTAGTAATAAGCCTTC 740
Db |||
Qy 850 TGAAGGACTTCACGTGACCTTTGCTGCTGAGATTTCAATACCTTAGTAATAAGCCTTC 909
Qy 741 AGAAGAAAGGTCAGGAGATCATTTGATGCTGCTCAAAATTCAGCAGAGTTTAAAC 800
Db |||
Qy 910 AGAAGAAAGGTCAGGAGATCATTTGATGCTGCTCAAAATTCAGCAGAGTTTAAAC 969
Qy 801 AGAAGCCTTCGACGTGGCCTCATTTGGAATGAATTTGATGAAACAGTACATTGA 860
Db |||
Qy 970 AGAAGCCTTCGACGTGGCCTCATTTGGAATGAATTTGATGAAACAGTACATTGA 1029
Qy 861 GTTTGTAGCTGACAGATTACTTTGGAACTTGGATTTCTCAAGGTTTTTCAGGCAGAAA 920
Db |||
Qy 1030 GTTTGTAGCTGACAGATTACTTTGGAACTTGGATTTCTCAAGGTTTTTCAGGCAGAAA 1089
Qy 921 TCCTTTTGTATGAGAAACATTTCTTTAGAGGAAAACAAATTTCTTTGAGAAACG 980
Db |||
Qy 1090 TCCTTTTGTATGAGAAACATTTCTTTAGAGGAAAACAAATTTCTTTGAGAAACG 1149
Qy 981 AGTTTCAGAGTATCAGCGTTTTCAGATTATGGCAGAAACACAGATAACGTTTCACCTT 1040
Db |||
Qy 1150 AGTTTCAGAGTATCAGCGTTTTCAGATTATGGCAGAAACACAGATAACGTTTCACCTT 1209
Qy 1041 GGATGCAGATTTT 1053
Db |||
Qy 1210 GGATGCAGATTTT 1222
Db |||
```

RESULT 5

```
AY398973
LOCUS AY398973
DEFINITION Homo sapiens HCM0069 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY398973
VERSION AY398973.1 GI:39754962
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1056)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
```

```
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
gene
ORIGIN
Query Match 89.7%; Score 945; DB 9; Length 1056;
Best Local Similarity 89.7%; Pred. No. 1.5e-240;
Matches 945; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
```

```
Qy 1 ATGGCGCACCGGAAAGCCGGAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 60
Db |||
Qy 61 GACACCAAGAAAGTGAATTAAGTCAATGAAGGCCACTCTTAGAAGAGTTCTCGC 120
Db |||
Qy 121 CGGTTTGTTCATCTTTCCAAATCCAGTACCTCGATATTGGAAAATGTAATAACAGGCACAG 180
Db |||
Qy 181 GCTTCTCTTCGGACAGCAGAGAGTTCGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
Qy 241 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCATCTTAGCCCTTTTTCGAGCCAGTGAT 300
Db |||
Qy 301 GGAATTGTAATGAAAATTTGGTGAGCGCTTTAGTCAGGAGTTCAGAGGCT 360
Db |||
Qy 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG 420
Db |||
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA 480
Db |||
Qy 481 ACCATGCCCTATGTTAAGAAAAGACAGATTTGGGCTTCGATGGATAGCAGATAGAAA 540
Db |||
Qy 541 TCTACTTTTGGGGAAGAGTGTGGCTTTTGTCTGTAGAGAGGAGTTTCTTCTCAGGA 600
Db |||
Qy 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Db |||
```

Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)

14671302
2 (bases 1 to 1056)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.

Location/Qualifiers
1..1056
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1_>1056
/locus_tag="HCM0069"

```
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTCGATGCATAGCAGATAGAAAA 540
Db 481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
Db 541 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 600
QY 601 TCTTTGCTGCTATATCTGCTAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
Db 601 TCAATTTGCTGCAATATCTGCTAAGAGAGAGGCTCATGCTGGACTGACTTTTTC 660
QY 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTCAA 720
Db 661 AATGAACATCATCAGCAGGAGTGAAGGACTTCACTGTGACTTTGCTGCTGATGTCAA 720
QY 721 TACTTTAGTAAATAAGCCTTCAGAGAGAGGCTTCAGGAGAGATCATTTGTGCTGCTCAA 780
Db 721 TACTTTAGTAAACAAGCCTTCAGAGATAGAGTGAAGGAAATCATTTGCTGCTGCTCAA 780
QY 781 ATTGAGCAGGAGTGTAAACAGAGAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTCGATT 840
Db 781 ATCGAG-----CAGAGCCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 827
QY 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTACTTTGGAACTTTGGATTCTCA 900
Db 828 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTACTTTGGAACTTTGGATTCTCA 887
QY 901 AAGGTTTTTTCAGGAGAGAAATCCTTTGATTTTATGGAACATTTCTTTAGAGGAAAA 960
Db 888 AAGATTTTTTCAGGAGAGAAATCCTTTGATTTTATGGAACATTTCTTTAGAGGAAAA 947
QY 961 ACAAAATTTCTTTGAGAAACAGTTTTCAGAGATACAGGCTTTTTCAGATTATGCGAGAAACC 1020
Db 948 ACAAAATTTCTTTGAGAAACAGTTTCTGAGATACAGGATTTGAGTATGCGCGAAACC 1007
QY 1021 ACAGATAACGCTTTCACCTTGGATGCAGATT 1052
Db 1008 ACAGATAATGCTTTCACCTTGGATGCAGATT 1039
```

```
RESULT 10
CN163214/c
LOCUS 952695 MARC 4PIG Sus scrofa cDNA 3', mRNA sequence. EST 02-APR-2004
ACCESSION CN163214
VERSION CN163214.1 GI:46177644
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 892)
Smith,T.P.L., Frenking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A.,
Nonnenman,D.J., Wray,J.E. and Keefe,J.W.
Porcine EST collection using a normalized library constructed from
embryos representing early developmental stages
Unpublished (2003)
JOURNAL Contact: Smith TPL
COMMENT USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross match v0.990329.
Plate: TW8060 row: G column: 4
Seq primer: TAGAAGCAGCAGTCGAG.
Location/Qualifiers
1. .892
/organism="Sus scrofa"
/mol_type="mRNA"
```

FEATURES
source

RESULT 11
CF995079
LOCUS

CF995079 914 bp mRNA linear EST 25-NOV-2003

/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 4PIG"
/note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
library made with combined RNA from day-10, day-13,
day-15, day-25, and day-30 whole embryos."

ORIGIN

```
Query Match 69.4%; Score 731.2; DB 7; Length 892;  
Best Local Similarity 94.1%; Pred. No. 1.4e-183;  
Matches 760; Conservative 0; Mismatches 48; Indels 0; Gaps 0;  
QY 246 AGCAGATGAGAAGTACTTCTCATCTCAGATCTTTAGTCAGAGGTGCAGGTTCCAGAGCTCGCTG 305  
Db 892 ATCAGATGAGAAGTATTTTATCTCTCAGATCTTTAGCTTTTTCAGCCAGTCATGGAAT 833  
QY 306 TGTAAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGCTCGCTG 365  
Db 832 TGTGAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGCGCTG 773  
QY 366 TTTCTATGCTTTCAAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTGTCTGAT 425  
Db 772 TTTCTACGCTTTCAAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTGTCTAAT 713  
QY 426 AGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTCGAAACCAT 485  
Db 712 AGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTCGAGACAT 653  
QY 486 GCCCTATGTTAAGAAAAGAGAGATTGGGCTTTGGATGGAATAGCAGATAGAGAAAATCTAC 545  
Db 652 GCCATATGTTAAGAAAAGAGAGATTGGGCTTTGAGATGGAATAGCAGATAGAGAAAATCTAC 593  
QY 546 TTTTGGGAAAAGAGTGGTGGCTTTGCTGCTAGAGAGATTCTTCTCAGGATCTTT 605  
Db 592 TTTTGGGAAAAGGTTGGTGGCTTTGCTGCTAGAGAGATTCTTCTCAGGATCTTT 533  
QY 606 TGCTGCTATATTTCTGGCTAAAAGAGAGAGGTCTTATGCGAGGACTCACATTTTCCAAATGA 665  
Db 532 TGCTGCTATATTTCTGGCTAAAAGAGAGGTCTTATGCGCTGCTCATGTTTCCAAATGA 473  
QY 666 ACTCATCAGCAGAGATGAAGGACTTTCATGTCATCTTGTGCTGCTGCTGCTGCTGCTGCT 725  
Db 472 ACTCATCAGCAGAGATGAAGGACTTTCATGTCATCTTGTGCTGCTGCTGCTGCTGCTGCT 413  
QY 726 AGTAAATGAAGCTTTCAGAAAAGAGGTTCAGGAGATCATTTGTGATGCTGTCAAATTTGA 785  
Db 412 GGTAAATGAAGCTTTCAGAAAAGAGGTTCAGGAGATCATTTGTGATGCTGTCAAATTTGA 353  
QY 786 GCAGAGTTTAAACAGAGAGCTTCCAGTTGGCTCATTTGGAATGAATTTGCAATTTGAT 845  
Db 352 GCAGAGTTTAAACAGAGAGCTTCCAGTTGGCTCATTTGGAATGAATTTGCAATTTGAT 293  
QY 846 GAAACAGTACATTCAGATTGTTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCAAAGGT 905  
Db 292 GAAACAGTATATTCAGATTGTTAGCTGACAGATTGCTTACAGAACTTGGATTCTCAAAGGT 233  
QY 906 TTTTCAGGAGAAAATCCTTTTGTATTTATGAAAACATTTCTTTAGAGGAAAACAA 965  
Db 232 TTTTCAGGAGAAAATCCTTTTGTATTTATGAAAACATTTCTTTAGAGGAAAACAA 173  
QY 966 TTTCTTTGAGAAACGAGTTTCAGAGTATCAGGTTTTCAGTATGCGGATGCGGAACTACAGA 1025  
Db 172 TTTCTTTGAGAAACGAGTTTCAGAGTATCAGGTTTTCAGTATGCGGATGCGGAACTACAGA 113  
QY 1026 TAAAGCTCTTACCTTGGATGCAGATTT 1053  
Db 112 TAAAGCTCTTACCTTGGATGCAGATTT 85
```

Swape, S.
10/698228
Seq. ID 12 w/Integ

> O <
O | | O IntelliGenetics
> O <

GENALIGN - Multiple Sequence Alignment Program
Release 5.4

Mon 26 Sep 105 11:01:11-PST

Solution Parameters:

Nucleic Alphabet = Identity
Output line length = 80
Compress = Off
Histogram = Off
Randomization = Off
AMINO-Res-length = 2
Deletion-weight = 5.00
Length-factor = 0
Matching-weight = 1.00
NUCLEIC-Res-length = 4
Spread-factor = 50

Clustered order of selected sequences:

- 1. US-10-698-228-2 (1-1053)
- 10. US-10-698-228-12 (1-1053)

-Region Alignment: (listed in clustered order)

```
US-10-698- 1 ATGGCGACCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCAG
US-10-698- 1 ATGGCGACCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCAG
consensus ATGGCGACCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCAG
US-10-698- 62 ACACCAAGAAAGTGAAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGCGG
US-10-698- 62 ACACCAAGAAAGTGAAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGCGG
consensus ACACCAAGAAAGTGAAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGCGG
US-10-698- 123 GTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAAATGTATAACAGGCAAGGCT
US-10-698- 123 GTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAAATGTATAACAGGCAAGGCT
consensus GTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAAATGTATAACAGGCAAGGCT
US-10-698- 184 TCCTTCTGGAAGAGAGAGGTGACTTATCAAAAGGATCTCCCTCACTGGAAACAGCTTA
US-10-698- 184 TCCTTCTGGAAGAGAGAGGTGACTTATCAAAAGGATCTCCCTCACTGGAAACAGCTTA
consensus TCCTTCTGGAAGAGAGAGGTGACTTATCAAAAGGATCTCCCTCACTGGAAACAGCTTA
US-10-698- 245 AAGCAGATGAGAGTACTCTCATCTCATCTTAGCCCTTTTTCAGCCAGTGGAT
US-10-698- 245 AAGCAGATGAGAGTACTCTCATCTCATCTTAGCCCTTTTTCAGCCAGTGGAT
consensus AAGCAGATGAGAGTACTCTCATCTCATCTTAGCCCTTTTTCAGCCAGTGGAT
US-10-698- 306 TGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAAGGTTCCAGAGGCTCGCTGT
US-10-698- 306 TGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAAGGTTCCAGAGGCTCGCTGT
consensus TGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAAGGTTCCAGAGGCTCGCTGT
```

```
US-10-698- 367 TTCTATGGCTTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGACAGTTTCTGCTGATAG
US-10-698- 367 TTCTATGGCTTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGACAGTTTCTGCTGATAG
consensus TTCTATGGCTTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGACAGTTTCTGCTGATAG
US-10-698- 428 ACACCTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAAACCATGCC
US-10-698- 428 ACACCTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAAACCATGCC
consensus ACACCTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAAACCATGCC
US-10-698- 489 CTATGTTAAGAAAAAGCAGATTGGGCCCTTGGCATGAGTAGCAGATAGAAATCTACTTTT
US-10-698- 489 CTATGTTAAGAAAAAGCAGATTGGGCCCTTGGCATGAGTAGCAGATAGAAATCTACTTTT
consensus CTATGTTAAGAAAAAGCAGATTGGGCCCTTGGCATGAGTAGCAGATAGAAATCTACTTTT
US-10-698- 550 GGGGAAAGAGTGGTGGCCCTTGGCTGCTAGAAAGGAGTTTCTTCTCAGGATCTTTTGGCTG
US-10-698- 550 GGGGAAAGAGTGGTGGCCCTTGGCTGCTAGAAAGGAGTTTCTTCTCAGGATCTTTTGGCTG
consensus GGGGAAAGAGTGGTGGCCCTTGGCTGCTAGAAAGGAGTTTCTTCTCAGGATCTTTTGGCTG
US-10-698- 611 CTATATTTCTGGCTAAAGAGAGAGGTCTTATGCGCAGGACTCACTTTTCCAAATGAACTCAT
US-10-698- 611 CTATATTTCTGGCTAAAGAGAGAGGTCTTATGCGCAGGACTCACTTTTCCAAATGAACTCAT
consensus CTATATTTCTGGCTAAAGAGAGAGGTCTTATGCGCAGGACTCACTTTTCCAAATGAACTCAT
US-10-698- 672 CAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGCTGATGTTTCCAAATCTTAGTAAAT
US-10-698- 672 CAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGCTGATGTTTCCAAATCTTAGTAAAT
consensus CAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGCTGATGTTTCCAAATCTTAGTAAAT
US-10-698- 733 AAGCCTTTCAGAAAGAGAGGTGAGGAGATCATTTGTTGATGCTGTCAAAATTTGAGCAGGAGT
US-10-698- 733 AAGCCTTTCAGAAAGAGAGGTGAGGAGATCATTTGTTGATGCTGTCAAAATTTGAGCAGGAGT
consensus AAGCCTTTCAGAAAGAGAGGTGAGGAGATCATTTGTTGATGCTGTCAAAATTTGAGCAGGAGT
US-10-698- 794 TTTTAAACAGAGCTTTGCCAGTTGGCCTCATTTGGAATGAATTTGATGAAACAGTGA
US-10-698- 794 TTTTAAACAGAGCTTTGCCAGTTGGCCTCATTTGGAATGAATTTGATGAAACAGTGA
consensus TTTTAAACAGAGCTTTGCCAGTTGGCCTCATTTGGAATGAATTTGATGAAACAGTGA
US-10-698- 855 CATTCAGTTTGTAGCTGACAGATTACTTGTGGAACTTTGGATTCTCAAGGTTTTTTCAGGCA
US-10-698- 855 CATTCAGTTTGTAGCTGACAGATTACTTGTGGAACTTTGGATTCTCAAGGTTTTTTCAGGCA
consensus CATTCAGTTTGTAGCTGACAGATTACTTGTGGAACTTTGGATTCTCAAGGTTTTTTCAGGCA
US-10-698- 916 GAAAAATCCTTTTGAATTTTATGAAAAACATTTCTTTAGAAAGAAAAACAAATTTCTTTGAGA
US-10-698- 916 GAAAAATCCTTTTGAATTTTATGAAAAACATTTCTTTAGAAAGAAAAACAAATTTCTTTGAGA
consensus GAAAAATCCTTTTGAATTTTATGAAAAACATTTCTTTAGAAAGAAAAACAAATTTCTTTGAGA
US-10-698- 977 AACGAGTTTTCAGAGTATCAGCGCTTTTGGCAGTTATGTCAGAGAAACCAAGATAAAGCTTTTCAC
```

```
US-10-698- 977 AACGAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGAGAAACACAGATAACGTCCTTCAC
consensus   AACGAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGAGAAACACAGATAACGTCCTTCAC

US-10-698- 1038 CTTGGATGCAGATTTT
              |||||
US-10-698- 1038 CTTGGATGCAGATTTT
consensus    CTTGGATGCAGATTTT
```

Alignment score = 1052.00
Scoring matrix:

	1	10
1	1052	
10		

TACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTGAAACCATGCCCTATGTTAAGAAAAA
|||||
TACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTGAAACCATGCCCTATGTTAAGAAAAA
440 450 460 470 480 490 500
510 520 530 540 550 560 570
GCAGATTGGGCTTTGGCATGGATAGCAGATAGAAAAATCTATTTTGGGAAAGAGTGGTGGCCTTTGCTGCT
|||||
GCAGATTGGGCTTTGGCATGGATAGCAGATAGAAAAATCTATTTTGGGAAAGAGTGGTGGCCTTTGCTGCT
510 520 530 540 550 560 570
580 590 600 610 620 630 640
GTAGAAGGAGTTTCTTCTCAGGATCTTTTGTCTGTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGA
|||||
GTAGAAGGAGTTTCTTCTCAGGATCTTTTGTCTGTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGA
580 590 600 610 620 630 640
650 660 670 680 690 700 710 720
CTCACTTTTCCAAATGAATCATCAGCAGAGATGAAGACTTCTACTGTGACTTTGCTTGCCTGATGTTCCAA
|||||
CTCACTTTTCCAAATGAATCATCAGCAGAGATGAAGACTTCTACTGTGACTTTGCTTGCCTGATGTTCCAA
650 660 670 680 690 700 710 720
730 740 750 760 770 780 790
TACTTAGTAATAAGCCTTCAGAGAGAGGCTCAGGGAGATCATTTGTGATGCTGTCAAAATTCAGCAGGAG
|||||
TACTTAGTAATAAGCCTTCAGAGAGAGGCTCAGGGAGATCATTTGTGATGCTGTCAAAATTCAGCAGGAG
730 740 750 760 770 780 790
800 810 820 830 840 850 860
TTTTTAACAGAGCCTTGGCAGTTGGCCTCATTTGGAATGAATTCATTTGATGAACAGTACATTCAGTTT
|||||
TTTTTAACAGAGCCTTGGCAGTTGGCCTCATTTGGAATGAATTCATTTGATGAACAGTACATTCAGTTT
800 810 820 830 840 850 860
870 880 890 900 910 920 930
GTAGCTCAGAGATTACTTGTGAACTTGGATTCTCAAAGGTTTTCAGGGCAGAAAAATCCTTTGATTTATG
|||||
GTAGCTCAGAGATTACTTGTGAACTTGGATTCTCAAAGGTTTTCAGGGCAGAAAAATCCTTTGATTTATG
870 880 890 900 910 920 930
940 950 960 970 980 990 1000
GAAAAATTTCTTTAGAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTT
|||||
GAAAAATTTCTTTAGAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTT
940 950 960 970 980 990 1000
1010 1020 1030 1040 1050 X
ATGCCAGAAACACAGATACGCTTTCACCTTGGATGCAGATTTT
|||||
ATGCCAGAAACACAGATACGCTTTCACCTTGGATGCAGATTTT
1010 1020 1030 1040 1050 X

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2005, 21:42:15 ; Search time 3258 Seconds
(without alignments)
15660.940 Million cell updates/sec

Title: US-10-698-228-12
Perfect score: 1053
Sequence: 1 atggcgacccggaaggcc.....tcacctggatgcagatttt 1053

Scoring table: IDENTITY_NUC
Gapo 10.0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: GenEmbl.*

- 1: gb_ba.*
- 2: gb_hsg.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1053	100.0	1053	6	AR454877 Sequence
2	1053	100.0	1053	6	BD064773 Novel pro
3	1053	100.0	1053	6	BD093086 Novel pro
4	1053	100.0	1081	6	AR454870 Sequence
5	1053	100.0	1081	6	BD064766 Novel pro
6	1053	100.0	1081	6	BD093079 Novel pro
7	1051.4	99.8	1053	6	AR454868 Sequence
8	1051.4	99.8	1053	6	BD064764 Novel pro
9	1051.4	99.8	1053	6	BD093077 Novel pro
10	1051.4	99.8	1056	6	CQ714252 Sequence
11	1051.4	99.8	1056	9	AB166671 Homo sapi
12	1051.4	99.8	4955	6	AR454869 Sequence
13	1051.4	99.8	4955	6	BD064765 Novel pro
14	1051.4	99.8	4955	6	BD093078 Novel pro
15	1051.4	99.8	4955	9	AB036063 Homo sapi
16	1048.2	99.5	1601	6	BD156916 Primer fo
17	1048.2	99.5	1601	6	AX877905 Sequence
18	1048.2	99.5	1601	9	AK001965 Homo sapi
19	896.8	85.2	4532	10	BC058103 Mus muscu

20	853.4	81.0	4571	9	BC042468 Homo sapi
21	851.8	80.9	900	9	AB163437 Homo sapi
22	839.2	79.7	171737	10	AC122379 Mus muscu
23	580.6	55.1	2500	6	AR300103 Sequence
24	580.6	55.1	2500	6	AR438503 Sequence
25	580.6	55.1	2500	6	AR562011 Sequence
26	580	55.1	1794	9	S40301 Homo sapien
27	580	55.1	2500	6	AR380825 Sequence
28	580	55.1	2500	6	AR333242 Sequence
29	580	55.1	2500	6	AX333687 Sequence
30	580	55.1	2500	6	AX334345 Sequence
31	580	55.1	2500	9	HSRR2SS X59618 H.sapiens R
32	578.4	54.9	1649	9	BC001886 Homo sapi
33	578.4	54.9	1653	9	BC030154 Homo sapi
34	578.4	54.9	3039	9	AK092671 Homo sapi
35	578.2	54.9	1562	5	BC061353 Xenopus t
36	573.6	54.5	217774	2	AC120603 Rattus no
37	573.2	54.4	3203	9	AK123010 Homo sapi
38	573	54.4	1538	5	CR762018 Xenopus t
39	573	54.4	1611	5	CR848103 Xenopus t
40	573	54.4	1647	5	BC080161 Xenopus t
41	570.4	54.2	2111	10	MUSRNRM2A M14223 Mouse ribon
42	568.8	54.0	1170	6	E14825 cDNA encodi
43	566.2	53.8	1570	5	BC072071 Xenopus l
44	564.2	53.6	2057	5	BC044355 Danio rer
45	562.6	53.4	1328	5	DRU57965 U57965 Danio rerio

ALIGNMENTS

RESULT 1
AR454877
LOCUS AR454877 1053 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 12 from patent US 6682917.
ACCESSION AR454877
VERSION AR454877.1 GI:42688832
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1053)
AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
TITILE Protein having a ribonucleotide Reductase activity and a DNA thereof
JOURNAL Patent: US 6682917-A 12 27-JAN-2004;
FEATURES Location/Qualifiers
1..1053
/organism="unknown"
/mol_type="genomic DNA"

QY	1	ATGGGCGACCCGGAAGCCGGAAGCGCGGCTGGATCAGGATGAGATCATCTTCA	60
Db	1	ATGGGCGACCCGGAAGCCGGAAGCGCGGCTGGATCAGGATGAGATCATCTTCA	60
QY	61	GACACCAAGCAAGTGAATAAAGTCAATGAAGGCCACTCTCTAAGAAAGAGTTCTCGC	120
Db	61	GACACCAAGCAAGTGAATAAAGTCAATGAAGGCCACTCTCTAAGAAAGAGTTCTCGC	120
QY	121	CGGTTTGTCATCTTTCCCAATCCAGTACCCCTGATATTGGAAATGTATAACAGGCACAG	180
Db	121	CGGTTTGTCATCTTTCCCAATCCAGTACCCCTGATATTGGAAATGTATAACAGGCACAG	180
QY	181	GCTTCCTTCTGGACAGCAGAAGAGTTGACTTATCAAGAGGATCTCCCTCACTGGAAACAAG	240
Db	181	GCTTCCTTCTGGACAGCAGAAGAGTTGACTTATCAAGAGGATCTCCCTCACTGGAAACAAG	240
QY	241	CTTAAAGCAGATGAGAAGTACTTCTCATCTCTACATCTTAGCCCTTTTGGAGCCAGTGAT	300

Db	241	CTTAAAGCAGATGAGAGTACTTCTCATCTCACACTTAGCCCTTTTTCAGCCAGTGAT	300
Qy	301	GGAAATGTAATGAAAAATTTGGTGAGCGCTTTAGTCAGAGGTGCGAGTTCCAGAGGCT	360
Db	301	GGAAATGTAATGAAAAATTTGGTGAGCGCTTTAGTCAGAGGTGCGAGTTCCAGAGGCT	360
Qy	361	CGCTGTTCTATGGCTTTCATAATCTCATCGAAGTGTTCACATCAGAGTGACAGTTTG	420
Db	361	CGCTGTTCTATGGCTTTCATAATCTCATCGAAGTGTTCACATCAGAGTGACAGTTTG	420
Qy	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAATTTTATTAATCAATTTGAA	480
Db	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAATTTTATTAATCAATTTGAA	480
Qy	481	ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCTTTCGATGATAGCAGATAGAAA	540
Db	481	ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCTTTCGATGATAGCAGATAGAAA	540
Qy	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA	600
Db	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA	600
Qy	601	TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC	660
Db	601	TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC	660
Qy	661	AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGGCTGTGATTTCCAA	720
Db	661	AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGGCTGTGATTTCCAA	720
Qy	721	TACTTAGTAAATAAGCCTTCAGAAAGAGGTTCAGGAGATCATTTGTTGATGCTCTCAA	780
Db	721	TACTTAGTAAATAAGCCTTCAGAAAGAGGTTCAGGAGATCATTTGTTGATGCTCTCAA	780
Qy	781	ATTGAGCAGGAGTCTTAAACAGAACCTTGGCCAGTTGGCTCATTTGGAATGAATTCATT	840
Db	781	ATTGAGCAGGAGTCTTAAACAGAACCTTGGCCAGTTGGCTCATTTGGAATGAATTCATT	840
Qy	841	TGATGAACAGTACATTGAGTTTGTAGCTGACAGATTAATTTGGAATTCATCA	900
Db	841	TGATGAACAGTACATTGAGTTTGTAGCTGACAGATTAATTTGGAATTCATCA	900
Qy	901	AAGGTTTTTCAGGCGAAGATCCTTTGATTTTATGGAACATTTCTTTAGAGGAAA	960
Db	901	AAGGTTTTTCAGGCGAAGATCCTTTGATTTTATGGAACATTTCTTTAGAGGAAA	960
Qy	961	ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTTATGCGAGAAC	1020
Db	961	ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTTATGCGAGAAC	1020
Qy	1021	ACAGATAACGCTTCCACCTTGGATGAGATTTT	1053
Db	1021	ACAGATAACGCTTCCACCTTGGATGAGATTTT	1053

RESULT 2
BD064773
LOCUS
DEFINITION Novel protein and DNA thereof.
ACCESSION BD064773
VERSION BD064773.1 GI:22610376
KEYWORDS JP 2001269184-A/10.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1053)
REFERENCE
AUTHORS Nakamura, Y., Arakawa, H. and Tanaka, H.
TITLE Novel protein and DNA thereof
JOURNAL Patent: JP 2001269184-A 10 02-OCT-2001;
YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD
COMMENT OS Homo sapiens (human)

PN	JP	2001269184-A/10	
PD	02-OCT-2001		
PI	27-JUN-2000	JP 2000192401	
PF	YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC		
C12N15/09	A61K31/711	A61K38/00, A61K38/43, A61K45/00, A61K48/00, PC	
A61P35/00,			
PC	A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC		
C12N5/10, C12N9/02,			
PC	C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53// (C12N15/09,		
PC	C12R1:191),		
PC	(C12N15/09, C12R1:19), (C12N9/02, C12R1:91), (C12P21/02, C12R1:91),		
PC	(C12P21/08, C12R1:91), (C12N15/00, A61K37/02, A61K37/48, C12N5/00,		
PC	(C12N15/00, C12R1:91), (C12N15/00, C12R1:19)		
CC	Novel protein and DNA thereof		
PH	Key	Location/Qualifiers	
FT	source	1..1053	
FT		/organism='Homo sapiens (human)'. Location/Qualifiers 1..1053 /organism='Homo sapiens' /mol_type='genomic DNA' /db_xref='taxon:9606'	
FEATURES			
Source			
ORIGIN			
Query Match 100.0%; Score 1053; DB 6; Length 1053;			
Best Local Similarity 100.0%; Pred. No. 6.4e-255; Indels 0; Gaps 0;			
Matches 1053; Conservative 0; Mismatches 0;			
Qy	1	ATGGGCGACCCGGAAGGCGGAGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA	60
Db	1	ATGGGCGACCCGGAAGGCGGAGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA	60
Qy	61	GCACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC	120
Db	61	GCACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC	120
Qy	121	CGGTTTGTCTATCTTCCCAATCCAGTACCTGATTTTGGAAATGTATTAACAGGACAG	180
Db	121	CGGTTTGTCTATCTTCCCAATCCAGTACCTGATTTTGGAAATGTATTAACAGGACAG	180
Qy	181	GCTTCCTTCTGGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAG	240
Db	181	GCTTCCTTCTGGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAG	240
Qy	241	CTTAAAGCAGATGAGAGTACTTCTCATCTTCAATCTTAGCCCTTTTTCAGCCAGTGAT	300
Db	241	CTTAAAGCAGATGAGAGTACTTCTCATCTTCAATCTTAGCCCTTTTTCAGCCAGTGAT	300
Qy	301	GGAAATGTAATGAAAAATTTGGTGAGCGCTTTAGTCAGAGGTGCGAGTTCCAGAGGCT	360
Db	301	GGAAATGTAATGAAAAATTTGGTGAGCGCTTTAGTCAGAGGTGCGAGTTCCAGAGGCT	360
Qy	361	CGCTGTTCTATGGCTTTCATAATCTCATCGAAGTGTTCACATCAGAGTGACAGTTTG	420
Db	361	CGCTGTTCTATGGCTTTCATAATCTCATCGAAGTGTTCACATCAGAGTGACAGTTTG	420
Qy	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAATTTTATTAATCAATTTGAA	480
Db	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAATTTTATTAATCAATTTGAA	480
Qy	481	ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCTTTCGATGATAGCAGATAGAAA	540
Db	481	ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCTTTCGATGATAGCAGATAGAAA	540
Qy	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA	600
Db	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA	600
Qy	601	TCTTTTGTGCTATATTTCTGGCTAAAGAGAGGCTTATTAAGCAGGACTCACTTTTCC	660
Db	601	TCTTTTGTGCTATATTTCTGGCTAAAGAGAGGCTTATTAAGCAGGACTCACTTTTCC	660
Qy	661	AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGGCTGTGATTTCCAA	720

Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTCCTGCTGATGTTCCAA 720
Qy 721 TACTTAGTAATAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGATGCTGTCAA 780
Db 721 TACTTAGTAATAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGATGCTGTCAA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCTTCCAGTTGGCCTCATTTGGAATGAATTCGATT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGCCTTCCAGTTGGCCTCATTTGGAATGAATTCGATT 840
Qy 841 TTGATGAACAGTACATTCAGTTGTAGCTGACAGATTAATTTGTTGGAACCTTGGATTTCA 900
Db 841 TTGATGAACAGTACATTCAGTTGTAGCTGACAGATTAATTTGTTGGAACCTTGGATTTCA 900
Qy 901 AAGTTTTTCAGGCGAAGAAATCCTTTGATTTTATGGAACATTTCTTTAGAGGAAAA 960
Db 901 AAGTTTTTCAGGCGAAGAAATCCTTTGATTTTATGGAACATTTCTTTAGAGGAAAA 960
Qy 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
Qy 1021 ACAGATAACGCTTCCACCTTGGATGCAGATTTT 1053
Db 1021 ACAGATAACGCTTCCACCTTGGATGCAGATTTT 1053

RESULT 3

BD093086

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

HIROSHI TANAKA

OS Homo sapiens (human)

PN WO 0100799-A/10

PD 04-JAN-2001

PF 27-JUN-2000 WO 2000JP04189

PR 28-JUN-1999 JP 99P 181131,06-JUL-1999 JP 99P 192391 PR

PI YUSUKE NAKAMURA,HIROFUMI ARAKAWA,HIROSHI TANAKA PC

C12N9/04,C12N15/53,C12N1/15,C12N1/19,C12N1/21,C12N5/10 PC

C12P21/02,A61K38/44,

PC A61K48/00,C07K16/40,G01N33/53,C12Q1/26

CC

FH

Key

Location/Qualifiers

1. .1053

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

FEATURES

source

ORIGIN

Query Match

Best Local Similarity

Matches 1053; Conservative

100.0%; Score 1053; DB 6; Length 1053;

Pred. No. 6 4e-255;

Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGACCGGAAAGCGGAGCGGCTGGATCAGGATGAGATCATCTTCA 60

Db 1 ATGGCGACCGGAAAGCGGAGCGGCTGGATCAGGATGAGATCATCTTCA 60

Qy 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGGCCACTCTTAAGAAAGAGTTCTCGC 120

Db 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGGCCACTCTCTAAGAAAGAGTTCTCGC 120
Qy 121 CGGTTTGTCTATCTTTCCAAATCCAGTACCTCTGATATTTGGAAAATGTATAAACAGGCACAG 180
Db 121 CGGTTTGTCTATCTTTCCAAATCCAGTACCTCTGATATTTGGAAAATGTATAAACAGGCACAG 180
Qy 181 GCTTCTCTCTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
Db 181 GCTTCTCTCTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
Qy 241 CTTTAAAGCAGATGAGAAGTACTTCTCTCATCTTACCTTTTGGAGCCAGTGAT 300
Db 241 CTTTAAAGCAGATGAGAAGTACTTCTCTCATCTTACCTTTTGGAGCCAGTGAT 300
Qy 301 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCCAGAGGCT 360
Db 301 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCCAGAGGCT 360
Qy 361 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTG 420
Db 361 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTTTAAATGCAATTTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTTTAAATGCAATTTGAA 480
Qy 481 ACCATGCCCTATGTTTAAAGAAAAAGCAGATTTGGCCTTGGATGGATAGAGATAGAAAA 540
Db 481 ACCATGCCCTATGTTTAAAGAAAAAGCAGATTTGGCCTTGGATGGATAGAGATAGAAAA 540
Qy 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGTTTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGTTTTTCTTCTCAGGA 600
Qy 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGTTCTTATGCCAGACACTCCTTTTTC 660
Db 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGTTCTTATGCCAGACACTCCTTTTTC 660
Qy 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTCTTGGCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTCTTGGCTGATGTTCCAA 720
Qy 721 TACTTAGTAATTAAGCCTTCAGAGAAAGGCTCAGAGAGATCATTTGTTGATGCTGTCAA 780
Db 721 TACTTAGTAATTAAGCCTTCAGAGAAAGGCTCAGAGAGATCATTTGTTGATGCTGTCAA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCTTCCAGAGGCTTGGCCTCATTTGGAATGAATTCGATT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGCCTTCCAGAGGCTTGGCCTCATTTGGAATGAATTCGATT 840
Qy 841 TTGATGAACAGTACATTCAGTTGTAGCTGACAGATTAATTTGTTGGAACCTTGGATTTCA 900
Db 841 TTGATGAACAGTACATTCAGTTGTAGCTGACAGATTAATTTGTTGGAACCTTGGATTTCA 900
Qy 901 AAGTTTTTCAGGCGAAGAAATCCTTTGATTTTATGGAACATTTCTTTAGAGGAAAA 960
Db 901 AAGTTTTTCAGGCGAAGAAATCCTTTGATTTTATGGAACATTTCTTTAGAGGAAAA 960
Qy 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGTTATGGCAGAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGTTATGGCAGAAACC 1020
Qy 1021 ACAGATAACGCTTCCACCTTGGATGCAGATTTT 1053
Db 1021 ACAGATAACGCTTCCACCTTGGATGCAGATTTT 1053

RESULT 4

AR454870

LOCUS

DEFINITION

ACCESSION

AR454870

Sequence 4 from patent US 6682917.

PAT 20-FEB-2004

AR454870

AR454870

QY	121	CGGTTTGTGTCATCTTTCCAAATCAGTACCTCGATATTTGGAAAATGTATAAACAGGCACAG	180
Db	140	CGGTTTGTGTCATCTTTCCAAATCAGTACCTCGATATTTGGAAAATGTATAAACAGGCACAG	199
QY	181	GCTTCCTTCTGACAGCAGAGAGGTGTGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG	240
Db	200	GCTTCCTTCTGACAGCAGAGAGGTGTGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG	259
QY	241	CTTAAAGCAGATGAGAGTACTTCTCATCTCTCATCTTAGCCCTTTTTCAGCCAGTGTAT	300
Db	260	CTTAAAGCAGATGAGAGTACTTCTCATCTCTCATCTTAGCCCTTTTTCAGCCAGTGTAT	319
QY	301	GGAAATGTAAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGTTCACAGGCT	360
Db	320	GGAAATGTAAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGTTCACAGGCT	379
QY	361	CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAAGATGTTTCACTCAGAGATGTACAGTTTG	420
Db	380	CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAAGATGTTTCACTCAGAGATGTACAGTTTG	439
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA	480
Db	440	CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA	499
QY	481	ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCCCTTGGATGGATAGCAGATAGAAAA	540
Db	500	ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCCCTTGGATGGATAGCAGATAGAAAA	559
QY	541	TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGTGTAGAAAGAGTTCCTTCTCAGGA	600
Db	560	TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGTGTAGAAAGAGTTCCTTCTCAGGA	619
QY	601	TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC	660
Db	620	TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC	679
QY	661	AATGAACCTCATCAGCAGAGATCAAGGACTTCACTGTGACTTTTGGCTTGCTGTATTTCCAA	720
Db	680	AATGAACCTCATCAGCAGAGATCAAGGACTTCACTGTGACTTTTGGCTTGCTGTATTTCCAA	739
QY	721	TACTTTAGTAAATAAGCCTTCAGAAAGAGGGTCAGGAGATCATTTGATGCTGTCAA	780
Db	740	TACTTTAGTAAATAAGCCTTCAGAAAGAGGGTCAGGAGATCATTTGATGCTGTCAA	799
QY	781	ATTGAGCAGGAGTTTAAACAGAGCCTTGCCAGTTCAGTTCAGGACTCACTTTTTC	840
Db	800	ATTGAGCAGGAGTTTAAACAGAGCCTTGCCAGTTCAGTTCAGGACTCACTTTTTC	859
QY	841	TTGATGAACACAGTACATTGAGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATTCTCA	900
Db	860	TTGATGAACACAGTACATTGAGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATTCTCA	919
QY	901	AAAGTTTTCAGGCAGAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGGAAAA	960
Db	920	AAAGTTTTCAGGCAGAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGGAAAA	979
QY	961	ACAAATTTCTTTAGAAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC	1020
Db	980	ACAAATTTCTTTAGAAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC	1039
QY	1021	ACAGATAAGCTTTTACCTTGGATGCAAGATTTT	1053
Db	1040	ACAGATAAGCTTTTACCTTGGATGCAAGATTTT	1072
RESULT 6			
BD093079			
LOCUS	BD093079	1081 bp	DNA linear PAT 27-AUG-2002
DEFINITION	Novel protein and its DNA.		
ACCESSION	BD093079		
VERSION	BD093079.1 GI:22638667		
KEYWORDS	WO 0100799-A/3		
SOURCE	Homo sapiens (human)		

ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 1081)		
AUTHORS	Nakamura,Y., Arakawa,H. and Tanaka,H.		
TITLE	Novel protein and its DNA		
JOURNAL	Patent: WO 0100799-A 3 04-JAN-2001; TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA		
COMMENT	OS Homo sapiens (human) PN WO 0100799-A/3 PD 04-JAN-2001 PF 27-JUN-2000 WO 2000JP004189 PR 28-JUN-1999 JP 99P 181131,06-JUL-1999 JP 99P 192391 PR PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC C12N9/04,C12N15/53,C12N1/15,C12N1/19,C12N1/21,C12N5/10 PC ,C12P21/02,A61K38/44, PC A61K48/00,C07K16/40,G01N33/53,C12Q1/26 CC		
FEATURES	PH Key	Location/Qualifiers	Location/Qualifiers.
Source	1..1081	/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
ORIGIN	Query Match 100.0%; Score 1053; DB 6; Length 1081; Best Local Similarity 100.0%; Pred. No. 6.4e-255; Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	ATGGGGACCCGAAAGCGCGGCGCGGCTGGATCAGGATGAGAGATCATCTTCA	60
Db	20	ATGGGGACCCGAAAGCGCGGCGCGGCTGGATCAGGATGAGAGATCATCTTCA	79
QY	61	GACACCAAGCAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC	120
Db	80	GACACCAAGCAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC	139
QY	121	CGGTTTGTGTCATCTTTCCAAATCAGTACCTCGATTTGGAAAATGTATAAACAGGCACAG	180
Db	140	CGGTTTGTGTCATCTTTCCAAATCAGTACCTCGATTTGGAAAATGTATAAACAGGCACAG	199
QY	181	GCTTCCTTCTGACAGCAGAGAGGTGTGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG	240
Db	200	GCTTCCTTCTGACAGCAGAGAGGTGTGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG	259
QY	241	CTTAAAGCAGATGAGAGTACTTCTCATCTCTCATCTTAGCCCTTTTTCAGCCAGTGTAT	300
Db	260	CTTAAAGCAGATGAGAGTACTTCTCATCTCTCATCTTAGCCCTTTTTCAGCCAGTGTAT	319
QY	301	GGAAATGTAAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGTTCACAGGCT	360
Db	320	GGAAATGTAAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGTTCACAGGCT	379
QY	361	CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAAGATTTTCACCTCAGAGATGTACAGTTTG	420
Db	380	CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAAGATTTTCACCTCAGAGATGTACAGTTTG	439
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA	480
Db	440	CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA	499
QY	481	ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCCCTTGGATGGATAGCAGATAGAAAA	540
Db	500	ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCCCTTGGATGGATAGCAGATAGAAAA	559
QY	541	TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGTGTAGAAAGAGTTCCTTCTCAGGA	600
Db	560	TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGTGTAGAAAGAGTTCCTTCTCAGGA	619
QY	601	TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC	660

```
Db      620 TCTTTTGCTGCTATATTCTGGCTAAAGAGAGAGGCTCTATGCCAGGACTCATCTTTTCC 679
Qy      661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGAGCTTTGCTTGCTGATGTTCCAA 720
Db      680 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGAGCTTTGCTTGCTGATGTTCCAA 739
Qy      721 TACTTAGTAAATAGCCCTTCAGAGAAAGGGTCAGGGAGATCAATTTGTTGATGCTGTCMAA 780
Db      740 TACTTAGTAAATAGCCCTTCAGAGAAAGGGTCAGGGAGATCAATTTGTTGATGCTGTCMAA 799
Qy      781 ATTGACGAGGAGTTTAAACAGAGCCCTTGCCAGTTGGCCTCATTTGGAATGAATTTGCATT 840
Db      800 ATTGACGAGGAGTTTAAACAGAGCCCTTGCCAGTTGGCCTCATTTGGAATGAATTTGCATT 859
Qy      841 TTGATGAACAGTACATTTGAGTTTCTAGCTGACAGATTACTTGTGGAACCTTGGATTTCTCA 900
Db      860 TTGATGAACAGTACATTTGAGTTTCTAGCTGACAGATTACTTGTGGAACCTTGGATTTCTCA 919
Qy      901 AAGGTTTTTTCAGGCGAGAAATCTTTTGTGATTTTATGGAACCAATTTCTTTAGAGGAAA 960
Db      920 AAGGTTTTTTCAGGCGAGAAATCTTTTGTGATTTTATGGAACCAATTTCTTTAGAGGAAA 979
Qy      961 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTGCAGTTATGGCAGAAACC 1020
Db      980 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTGCAGTTATGGCAGAAACC 1039
Qy      1021 ACAGATACGCTTTCACCTTGGATCAGAGTTT 1053
Db      1040 ACAGATAACGCTTTCACCTTGGATCAGAGTTT 1072
```

```
RESULT 7
AR454868
LOCUS      AR454868      1053 bp      DNA      linear      PAT 20-FEB-2004
DEFINITION Sequence 2 from patent US 6682917.
ACCESSION AR454868
VERSION   AR454868.1 GI:42688823
KEYWORDS
SOURCE    Unknown.
ORGANISM  Unknown.
           Unclassified.
```

REFERENCE 1 (bases 1 to 1053)
Nakamura,Y., Arakawa,H. and Tanaka,H.
Protein having a ribonucleotide Reductase activity and a DNA

TITLE thereof
JOURNAL Patent: US 6682917-A 2 27-JAN-2004;
FEATURES Location/Qualifiers
source 1..1053
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

```
Query Match 99.8%; Score 1051.4; DB 6; Length 1053;
Best Local Similarity 99.9%; Pred. No. 1.6e-254;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 ATGGCGGACCCGGAAGCCGGAAGCCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db      1 ATGGCGGACCCGGAAGCCGGAAGCCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Qy      61 GACACCAACGAAAGTGAATAAAGTCAATGAAGGCCACTCTTAAGAAAGGTTCTCGC 120
Db      61 GACACCAACGAAAGTGAATAAAGTCAATGAAGGCCACTCTTAAGAAAGGTTCTCGC 120
Qy      121 CGGTTTGTCTATCTTTCCCAATCAGTACCTCTGATTTTGGAAAATGTATAAACAGGCACAG 180
Db      121 CGGTTTGTCTATCTTTCCCAATCAGTACCTCTGATTTTGGAAAATGTATAAACAGGCACAG 180
Qy      181 GCTTCTCTTCTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAAG 240
Db      181 GCTTCTCTTCTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAAG 240
```

```
Qy      241 CTTAAAGCAGATCAGAAAGTACTTCTCTCATCATCTTAGCCCTTTTTCGAGCCAGTGAT 300
Db      241 CTTAAAGCAGATCAGAAAGTACTTCTCTCTCATCATCTTAGCCCTTTTTCGAGCCAGTGAT 300
Qy      301 GGAAATGTAAATGAAATTTTGGTGGAGCGCTTTAGTCAGAGGTGCAAGGTTCCAGAGGCT 360
Db      301 GGAAATGTAAATGAAATTTTGGTGGAGCGCTTTAGTCAGAGGTGCAAGGTTCCAGAGGCT 360
Qy      361 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAAGTGTTCACTCAGAGATGTACAGTTG 420
Db      361 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAAGTGTTCACTCAGAGATGTACAGTTG 420
Qy      421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGAA 480
Db      421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGAA 480
Qy      481 ACATGCCCTTATGTTAAAGAAAAAGCAGATTTGGGCCCTTGCAGTGGATAGCAGATAGAAAA 540
Db      481 ACCATGCCCTTATGTTAAAGAAAAAGCAGATTTGGGCCCTTGCAGTGGATAGCAGATAGAAAA 540
Qy      541 TCTACTTTTGGGGAAGAGTGGTGGCCTTTTGTGCTGTAGAAGAGTTCCTTCTCTCAGGA 600
Db      541 TCTACTTTTGGGGAAGAGTGGTGGCCTTTTGTGCTGTAGAAGAGTTCCTTCTCTCAGGA 600
Qy      601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGTCCTTATGCGCAGGACTCAGCTTTTTC 660
Db      601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGTCCTTATGCGCAGGACTCAGCTTTTTC 660
Qy      661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db      661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Qy      721 TACTTAGTAAATAGCCCTTCAGAGAAAGGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db      721 TACTTAGTAAATAGCCCTTCAGAGAAAGGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Qy      781 ATTGAGCAGGAGTTTAAACAGAGCCCTTGCCAGTTGGCCTCATTTGGAATGAATTTGATT 840
Db      781 ATTGAGCAGGAGTTTAAACAGAGCCCTTGCCAGTTGGCCTCATTTGGAATGAATTTGATT 840
Qy      841 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTGTGGAATTTGATTTCTCA 900
Db      841 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTGTGGAATTTGATTTCTCA 900
Qy      901 AAGGTTTTTTCAGGCGAGAAATCTTTTGTGATTTTATGGAACCAATTTCTTTAGAGGAAA 960
Db      901 AAGGTTTTTTCAGGCGAGAAATCTTTTGTGATTTTATGGAACCAATTTCTTTAGAGGAAA 960
Qy      961 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTTCAGTTATGGCAGAAACC 1020
Db      961 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTTCAGTTATGGCAGAAACC 1020
Qy      1021 ACAGATAACGCTTTCACCTTGGATCAGAGTTT 1053
Db      1021 ACAGATAACGCTTTCACCTTGGATCAGAGTTT 1053

RESULT 8
BD064764
LOCUS      BD064764      1053 bp      DNA      linear      PAT 27-AUG-2002
DEFINITION Novel protein and DNA thereof.
ACCESSION BD064764
VERSION   BD064764.1 GI:22610367
KEYWORDS JP 2001269184-A/1.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1053)
AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE    Novel protein and DNA thereof
PATENT   JP 2001269184-A 1 02-OCT-2001;
JOURNAL YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD
```



```
COMMENT OS Homo sapiens (human)
PN: JP 2001269184-A/1
PD: 02-OCT-2001
PF 27-JUN-2000 JP 2000192401
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N15/09, A61K31/711, A61K38/00, A61K38/43, A61K45/00, A61K48/00, PC
A61P35/00,
PC A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC
, C12N5/10, C12N9/02,
PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53// (C12N15/09,
PC C12R1/91),
PC (C12N15/09, C12R1/19), (C12N9/02, C12R1/91), (C12P21/02, C12R1/91),
PC (C12P21/08, C12R1/91), (C12N15/00, A61K37/02, A61K37/48, C12N5/00,
PC (C12N15/00, C12R1/91), (C12N15/00, C12R1/19)
CC Novel protein and DNA thereof
FH Key Location/Qualifiers
FT source 1..1053
FT /organism='Homo sapiens (human)'.
FEATURES
source
1..1053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 1053;
Best Local Similarity 99.9%; Pred. No. 1.6e-254;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 ATGGCGACCGCGAAAGCGGAGCGCGCGGCTGGATCAGATCAGATCAGATCATCTTCA 60
Db 1 ATGGCGACCGCGAAAGCGGAGCGCGCGGCTGGATCAGATCAGATCAGATCATCTTCA 60
Qy 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Qy 121 CGGTTTGTCATCTTCCGATCCAGTACCTGATATTTGGAAATATTAACAGGACAG 180
Db 121 CGGTTTGTCATCTTCCGATCCAGTACCTGATATTTGGAAATATTAACAGGACAG 180
Qy 181 GCTTCCTTCTGGACAGCAGAGAGGTTGACTTATCAAGAGGATCTCCCTCACTGGAACAG 240
Db 181 GCTTCCTTCTGGACAGCAGAGAGGTTGACTTATCAAGAGGATCTCCCTCACTGGAACAG 240
Qy 241 CTTAAGCAGATGAGAAGTACTTCACTCTCTCACATCTTAGCCTTTTTCGAGCCAGTGAT 300
Db 241 CTTAAGCAGATGAGAAGTACTTCACTCTCTCACATCTTAGCCTTTTTCGAGCCAGTGAT 300
Qy 301 GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGAGGTCGAGGTTCCAGAGGCT 360
Db 301 GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGAGGTCGAGGTTCCAGAGGCT 360
Qy 361 CGCTGTTTCTATGCTTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGCTTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
Qy 481 ACCATGCCCTATGTTAGAAAAAGAGAGATGGGCTTGCATGATGATAGCAGATGAAAA 540
Db 481 ACCATGCCCTATGTTAGAAAAAGAGAGATGGGCTTGCATGATGATAGCAGATGAAAA 540
Qy 541 TCTACTTTTGGGAAAGAGTGGCTTTGCTGCTGAGAGAGGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGCTTTGCTGCTGAGAGAGGTTTCTTCTCAGGA 600
Qy 601 TCTTTTGCTGCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCATTCTTCC 660
Db 601 TCTTTTGCTGCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCATTCTTCC 660
Qy 661 AATGAACTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 661 AATGAACTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Qy 721 TACTTAGTAAATAAGCCTTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db 721 TACTTAGTAAATAAGCCTTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
Qy 781 ATTGAGCAGAGATTTTAAACAGAGGCTTGCAGGATGCGCTCATTTGAATGAATTCATT 840
Db 781 ATTGAGCAGAGATTTTAAACAGAGGCTTGCAGGATGCGCTCATTTGAATGAATTCATT 840
Qy 841 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATCTCA 900
Db 841 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATCTCA 900
Qy 901 AAGGTTTTTTCAGCAGAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGGAAAA 960
Db 901 AAGGTTTTTTCAGCAGAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGGAAAA 960
Qy 961 ACAAAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGGTTTTCAGGTTTTCAGGTTTTCAG 1020
Db 961 ACAAAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGGTTTTCAGGTTTTCAGGTTTTCAG 1020
Qy 1021 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1053
Db 1021 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1053
RESULT 9
BD093077 Novel protein and its DNA.
LOCUS BD093077 1053 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel protein and its DNA.
ACCESSION BD093077
VERSION BD093077.1 GI:22638665
KEYWORDS WO 0100799-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1053)
AUTHORS Nakamura, Y., Arakawa, H. and Tanaka, H.
TITLE Novel protein and its DNA
JOURNAL Patent: WO 0100799-A 1 04-JAN-2001;
TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,
HIROSHI TANAKA
COMMENT OS Homo sapiens (human)
PN WO 0100799-A/1
PD 04-JAN-2001
PF 27-JUN-2000 WO 2000JP004189
PR 28-JUN-1999 JP 99P 181131, 06-JUL-1999 JP 99P 192391 PR
21-JAN-2000 JP 00P 017770
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC
, C12P21/02, A61K38/44,
PC A61K48/00, C07K16/40, G01N33/53, C12Q1/26
CC
FH Key Location/Qualifiers.
FEATURES
source
1..1053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 1053;
Best Local Similarity 99.9%; Pred. No. 1.6e-254;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 ATGGCGACCGCGAAAGCGGAGCGCGCGGCTGGATCAGATCAGATCAGATCATCTTCA 60
Db 1 ATGGCGACCGCGAAAGCGGAGCGCGCGGCTGGATCAGATCAGATCAGATCATCTTCA 60
Qy 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Qy 121 CGGTTTGTCATCTTCCGATCCAGTACCTGATATTTGGAAATATTAACAGGACAG 180
Db 121 CGGTTTGTCATCTTCCGATCCAGTACCTGATATTTGGAAATATTAACAGGACAG 180
Qy 181 GCTTCCTTCTGGACAGCAGAGAGGTTGACTTATCAAGAGGATCTCCCTCACTGGAACAG 240
Db 181 GCTTCCTTCTGGACAGCAGAGAGGTTGACTTATCAAGAGGATCTCCCTCACTGGAACAG 240
Qy 241 CTTAAGCAGATGAGAAGTACTTCACTCTCTCACATCTTAGCCTTTTTCGAGCCAGTGAT 300
Db 241 CTTAAGCAGATGAGAAGTACTTCACTCTCTCACATCTTAGCCTTTTTCGAGCCAGTGAT 300
Qy 301 GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGAGGTCGAGGTTCCAGAGGCT 360
Db 301 GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGAGGTCGAGGTTCCAGAGGCT 360
Qy 361 CGCTGTTTCTATGCTTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGCTTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
Qy 481 ACCATGCCCTATGTTAGAAAAAGAGAGATGGGCTTGCATGATGATAGCAGATGAAAA 540
Db 481 ACCATGCCCTATGTTAGAAAAAGAGAGATGGGCTTGCATGATGATAGCAGATGAAAA 540
Qy 541 TCTACTTTTGGGAAAGAGTGGCTTTGCTGCTGAGAGAGGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGCTTTGCTGCTGAGAGAGGTTTCTTCTCAGGA 600
Qy 601 TCTTTTGCTGCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCATTCTTCC 660
Db 601 TCTTTTGCTGCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCATTCTTCC 660
```



```
QY 61 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 121 CGGTTTGTCACTTTCCAAATCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG 180
Db 121 CGGTTTGTCACTTTCCAAATCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG 180
QY 181 GCTTCTCTCGACAGCAGAGAGGTTGACATTATCAAGAGTCTCCCTCACTGGACCAAG 240
Db 181 GCTTCTCTCGACAGCAGAGAGGTTGACATTATCAAGAGTCTCCCTCACTGGACCAAG 240
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 300
QY 301 GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
Db 301 GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTTTCAATCTCATTCAGAAATGTTCACTCAGAGTGTACAGTTG 420
Db 361 CGCTGTTTCTATGGCTTTTCAATCTCATTCAGAAATGTTCACTCAGAGTGTACAGTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
QY 481 ACCATGCCCCTATGTTAAGAAAAAGCAGATTGGGCTTTGCGATGGATGACAGATAGAAA 540
Db 481 ACCATGCCCCTATGTTAAGAAAAAGCAGATTGGGCTTTGCGATGGATGACAGATAGAAA 540
QY 541 TCTACTTTTGGGAAAGAGTGGTGCCCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGCCCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
QY 601 TCTTTTGTCTATATTCGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Db 601 TCTTTTGTCTATATTCGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTTCGCTGTATGTTCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTTCGCTGTATGTTCAA 720
QY 721 TACTTAGTAAATAGCCCTTCAGAGAAAGGTCAGGAGATCAATTTGTAGCTGTCAA 780
Db 721 TACTTAGTAAATAGCCCTTCAGAGAAAGGTCAGGAGATCAATTTGTAGCTGTCAA 780
QY 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGGCAGTTGGCCTCATTTGGAATGAATTCAT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGGCAGTTGGCCTCATTTGGAATGAATTCAT 840
QY 841 TTGATGAACAGTACATTTAGTTTGTAGCTGACAGATTACTTTGTGGAATCTGATTTCTCA 900
Db 841 TTGATGAACAGTACATTTAGTTTGTAGCTGACAGATTACTTTGTGGAATCTGATTTCTCA 900
QY 901 AAGGTTTTTCAGGCGAAAAATCCCTTTGATTTTATGGAACCAATTTCTTAGAAGGAAA 960
Db 901 AAGGTTTTTCAGGCGAAAAATCCCTTTGATTTTATGGAACCAATTTCTTAGAAGGAAA 960
QY 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGSCAGAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGSCAGAAACC 1020
QY 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053
Db 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053
```

RESULT 10
CQ714252
LOCUS
DEFINITION Sequence 186 from Patent WO02068579.
1056 bp DNA linear PAT 03-FEB-2004

```
ACCESSION CQ714252
VERSION CQ714252.1 GI:42275109
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 186 06-SEP-2002;
PE Corporation (NV) (US)
FEATURES
source
location/Qualifiers
1..1056
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
```

ORIGIN

Query Match 99.8%; Score 1051.4; DB 6; Length 1056;
Best Local Similarity 99.9%; Pred. No. 1.6e-254;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGGGGGACCCGGAAGGCGGGAAGCGCGCGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGGGGACCCGGAAGGCGGGAAGCGCGCGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 121 CGGTTTGTCACTTTCCAAATCAGTACCTGATATTTGAAATATGATAAACAGGCACAG 180
Db 121 CGGTTTGTCACTTTCCAAATCAGTACCTGATATTTGAAATATGATAAACAGGCACAG 180
QY 181 GCTTCTCTCGACAGCAGAGGTTGACATTATCAAGAGTCTCCCTCACTGGAAACAG 240
Db 181 GCTTCTCTCGACAGCAGAGGTTGACATTATCAAGAGTCTCCCTCACTGGAAACAG 240
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCATCTTAGCCCTTTTTCAGCCAGTGAT 300
QY 301 GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
Db 301 GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTTTCAATCTCATTCAGAGATGTTCACTCAGAGTGTACAGTTG 420
Db 361 CGCTGTTTCTATGGCTTTTCAATCTCATTCAGAGATGTTCACTCAGAGTGTACAGTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
QY 481 ACCATGCCCCTATGTTAAGAAAAAGCAGATTGGGCTTTGCGATGGATGACAGATAGAAA 540
Db 481 ACCATGCCCCTATGTTAAGAAAAAGCAGATTGGGCTTTGCGATGGATGACAGATAGAAA 540
QY 541 TCTACTTTTGGGAAAGAGTGGTGCCCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGCCCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
QY 601 TCTTTTGTCTATATTCGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Db 601 TCTTTTGTCTATATTCGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTTCGCTGTATGTTCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTTCGCTGTATGTTCAA 720
```

QY 721 TACTAGTAATAAGCCCTCAGAGAAAGGGTCAGGAGATCAATGTTGATGCTCTCAA 780
Db 721 TACTAGTAATAAGCCCTCAGAGAAAGGGTCAGGAGATCAATGTTGATGCTCTCAA 780
QY 781 ATTGAGCAGGAGTTTAAACAGAACCTTGCAGTTGGCTCATTTGGAATGAATTCAT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAACCTTGCAGTTGGCTCATTTGGAATGAATTCAT 840
QY 841 TTGATGAACAGTACATTTGAGTTTCTAGCTGACAGATTACTTGTGGAACCTTGGATTTCTCA 900
Db 841 TTGATGAACAGTACATTTGAGTTTCTAGCTGACAGATTACTTGTGGAACCTTGGATTTCTCA 900
QY 901 AAGGTTTTTCAGGCAAGAAATCCCTTTTGGATTTTATGGAAGAAATTTCTTTAGAGGAAAA 960
Db 901 AAGGTTTTTCAGGCAAGAAATCCCTTTTGGATTTTATGGAAGAAATTTCTTTAGAGGAAAA 960
QY 961 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGATTTGCGAGAAACC 1020
Db 961 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGATTTGCGAGAAACC 1020
QY 1021 ACAGATACGCTTTCACCTTGGATGCGAGATTTT 1053
Db 1021 ACAGATACGCTTTCACCTTGGATGCGAGATTTT 1053

RESULT 11
AB166671
LOCUS
DEFINITION Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase small subunit 2, complete cds.
VERSION AB166671
KEYWORDS AB166671.1 GI:45259568
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ugai, H. and Yokoyama, K. K.
TITLE Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase small subunit 2, complete cds
JOURNAL Published Only in Database (2004)
REFERENCE 2 (bases 1 to 1056)
AUTHORS Ugai, H. and Yokoyama, K. K.
TITLE Direct Submission
JOURNAL Submitted (05-MAR-2004) Hideyo Ugai, RIKEN BioResource Center, Gene Engineering Division; 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan (E-mail: ugai@rtc.riken.go.jp, Tel: 81-29-836-3612, Fax: 81-29-836-9120)
FEATURES
source
1. .1056
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="p53R2.C19"
/cell_line="Saos-2"
/cell_type="Osteogenic sarcoma"
1. .1056
/gene="p53R2"
1. .1056
/gene="p53R2"
/codon_start=1
/product="p53-inducible ribonucleotide reductase small subunit 2"
/protein_id="BAD12267.1"
/db_xref="GI:45259569"
/translation="MGDPPEAAGLDQDSSDTHESBIKSNPEELLKSRFRVIFPIQDWMKQAQAFWTAEEVDLSKDLPHWNKLUKADKEYFISHLAFPAASDGI VNNELVRFQGVQVQPARCFYQFQILINVHSMYSLDITVIRDPKREFFLNALIE TNPYVKKADWALRDLAKRSTFGVRVFAAAGEVFISGSAFPAIPLWKKRGLMPGLT FSNELISRDGLHCDACLMOYLVNKPSRVERREIIVDAVKI BOEPLTALPVGLIG MNCILMKQYIEFVADRLVELGFSKVFQENPPDFMENISLEGKTNFEKRVSEYQRF AVNAETDNTVDLTDADE"

QY 721 TACTAGTAATAAGCCCTCAGAGAAAGGGTCAGGAGATCAATGTTGATGCTCTCAA 780
Db 721 TACTAGTAATAAGCCCTCAGAGAAAGGGTCAGGAGATCAATGTTGATGCTCTCAA 780
QY 781 ATTGAGCAGGAGTTTAAACAGAACCTTGCAGTTGGCTCATTTGGAATGAATTCAT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAACCTTGCAGTTGGCTCATTTGGAATGAATTCAT 840
QY 841 TTGATGAACAGTACATTTGAGTTTCTAGCTGACAGATTACTTGTGGAACCTTGGATTTCTCA 900
Db 841 TTGATGAACAGTACATTTGAGTTTCTAGCTGACAGATTACTTGTGGAACCTTGGATTTCTCA 900
QY 901 AAGGTTTTTCAGGCAAGAAATCCCTTTTGGATTTTATGGAAGAAATTTCTTTAGAGGAAAA 960
Db 901 AAGGTTTTTCAGGCAAGAAATCCCTTTTGGATTTTATGGAAGAAATTTCTTTAGAGGAAAA 960
QY 961 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGATTTGCGAGAAACC 1020
Db 961 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGATTTGCGAGAAACC 1020
QY 1021 ACAGATACGCTTTCACCTTGGATGCGAGATTTT 1053
Db 1021 ACAGATACGCTTTCACCTTGGATGCGAGATTTT 1053

Query Match 99.8%; Score 1051.4; DB 9; Length 1056;
Best Local Similarity 99.9%; Pred. No. 1.6e-254;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGGGCGGCGGCTCGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGGCGACCCGGAAGGGCGGCGGCTCGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
QY 121 CGGTTTGTGCTATCTTTCCAAATCAGTACCTGATATTTGGAAAATGTAATAACAGGCACAG 180
Db 121 CGGTTTGTGCTATCTTTCCAAATCAGTACCTGATATTTGGAAAATGTAATAACAGGCACAG 180
QY 181 GCTTCCTTCTTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db 181 GCTTCCTTCTTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
QY 241 CTTAAGCAGATGAGAAAGTACTTCTCTCATCTTACATCTTAGCCCTTTTTCAGAGCAGTGAT 300
Db 241 CTTAAGCAGATGAGAAAGTACTTCTCTCATCTTACATCTTAGCCCTTTTTCAGAGCAGTGAT 300
QY 301 GGAATTTGAATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGTTCAGAGGCT 360
Db 301 GGAATTTGAATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGTTCAGAGGCT 360
QY 361 GCGTGTCTTCTAGGCTTTCAAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTCAGTTG 420
Db 361 GCGTGTCTTCTAGGCTTTCAAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTCAGTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTAATCAATTTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTAATCAATTTGAA 480
QY 481 ACCATGCTCTATGTTAAGAAAAAGCAGATTGGGCTTCGGATGGATAGCAGATAGAAAA 540
Db 481 ACCATGCTCTATGTTAAGAAAAAGCAGATTGGGCTTCGGATGGATAGCAGATAGAAAA 540
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAAGAGGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAAGAGGAGTTTCTTCTCAGGA 600
QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGCTTTATGCGAGAGCTCACTTTTTC 660
Db 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGCTTTATGCGAGAGCTCACTTTTTC 660
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
QY 721 TACTTAGTAAATAGCCCTTCAGAAAGGGTCAGGAGATCATTTGATGCTGTCAAA 780
Db 721 TACTTAGTAAATAGCCCTTCAGAAAGGGTCAGGAGATCATTTGATGCTGTCAAA 780
QY 781 ATTGAGCAGGAGTTTAAACAGAACCTTGCAGTTGGCTCATTTGGAATGAATTCAT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAACCTTGCAGTTGGCTCATTTGGAATGAATTCAT 840
QY 841 TTGATGAACAGTACATTTGAGTTTCTAGCTGACAGATTACTTGTGGAACCTTGGATTTCTCA 900
Db 841 TTGATGAACAGTACATTTGAGTTTCTAGCTGACAGATTACTTGTGGAACCTTGGATTTCTCA 900
QY 901 AAGGTTTTTCAGGCAAGAAATCCCTTTTGGATTTTATGGAAGAAATTTCTTTAGAGGAAAA 960
Db 901 AAGGTTTTTCAGGCAAGAAATCCCTTTTGGATTTTATGGAAGAAATTTCTTTAGAGGAAAA 960
QY 961 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGATTTGCGAGAAACC 1020
Db 961 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGATTTGCGAGAAACC 1020

```
QY 1021 ACAGATACGCTTCCACCTTGATGACAGATTTT 1053
| | | | | | | | | | | | | | | | | | | | | |
Db 1021 ACAGATAACGCTTCCACCTTGATGACAGATTTT 1053

RESULT 12
AR454869
LOCUS AR454869 4955 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 3 from patent US 6682917.
ACCESSION AR454869
VERSION AR454869.1 GI:42688824
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 4955)
AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE Protein having a ribonucleotide Reductase activity and a DNA
thereof
JOURNAL Patent: US 6682917-A 3 27-JAN-2004;
FEATURES
source
1..4955
/mol_type="genomic DNA"

ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 4955;
Best Local Similarity 99.9%; Pred. No. 1.5e-254;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGGACCGGAAGCCGGAAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 60
| | | | | | | | | | | | | | | | | | | | | |
Db 245 ATGGCGGACCGCGGAAGCCGGAAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 304
| | | | | | | | | | | | | | | | | | | | | |

QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGACCACTCTTAAGAAAGAGTTCTCGC 120
| | | | | | | | | | | | | | | | | | | | | |
Db 305 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGACCACTCTTAAGAAAGAGTTCTCGC 364
| | | | | | | | | | | | | | | | | | | | | |

QY 121 CGGTTTGTCATCTTTCCAAATCAGTACCTGTATTTTGGAAATGTATAAACAGGCACAG 180
| | | | | | | | | | | | | | | | | | | | | |
Db 365 CGGTTTGTCATCTTTCCAAATCAGTACCTGTATTTTGGAAATGTATAAACAGGCACAG 424
| | | | | | | | | | | | | | | | | | | | | |

QY 181 GCTTCCTTCTGGACAGCAGAGAGTTGACTTATCAAGGATCTCCCTCACTGGACACAG 240
| | | | | | | | | | | | | | | | | | | | | |
Db 425 GCTTCCTTCTGGACAGCAGAGAGTTGACTTATCAAGGATCTCCCTCACTGGACACAG 484
| | | | | | | | | | | | | | | | | | | | | |

QY 241 CTTAAGCAGATGAGAAGTACTTCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT 300
| | | | | | | | | | | | | | | | | | | | | |
Db 485 CTTAAGCAGATGAGAAGTACTTCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT 544
| | | | | | | | | | | | | | | | | | | | | |

QY 301 GGAATTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
| | | | | | | | | | | | | | | | | | | | | |
Db 545 GGAATTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 604
| | | | | | | | | | | | | | | | | | | | | |

QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAATGTTCACTCAGAGATGTACAGTTTG 420
| | | | | | | | | | | | | | | | | | | | | |
Db 605 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG 664
| | | | | | | | | | | | | | | | | | | | | |

QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTAATCAATTGAA 480
| | | | | | | | | | | | | | | | | | | | | |
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTAATCAATTGAA 724
| | | | | | | | | | | | | | | | | | | | | |

QY 481 ACCATGCCCTATGTTTAAGAAAAAGCAGATTGGCGCTTTCGATGGATAGCAGATAGAAA 540
| | | | | | | | | | | | | | | | | | | | | |
Db 725 ACCATGCCCTATGTTTAAGAAAAAGCAGATTGGCGCTTTCGATGGATAGCAGATAGAAA 784
| | | | | | | | | | | | | | | | | | | | | |

QY 541 TCTACTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAAAGAGTTTCTTCTCAGGA 600
| | | | | | | | | | | | | | | | | | | | | |
Db 785 TCTACTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAAAGAGTTTCTTCTCAGGA 844
| | | | | | | | | | | | | | | | | | | | | |

QY 601 TCTTTTGTGCTATATCTGGCTTAAGAGAGAGGCTTATGCCAGGACTCATCTTTTCC 660
| | | | | | | | | | | | | | | | | | | | | |
```

Best Local Similarity 99.9%; Pred. No. 1.5e-254; Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	1	ATGGCGACCGGAAAGCGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA	60
Db	245	ATGGCGACCGGAAAGCGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA	304
QY	61	GACACCAACGAAAGTGAATAAGTCAAAATGAAGACCACTCCTAAGAAAGAGTTCTCGC	120
Db	305	GACACCAACGAAAGTGAATAAGTCAAAATGAAGACCACTCCTAAGAAAGAGTTCTCGC	364
QY	121	CGTTTGTCACTTTTCCAAATCCAGTACCTGTATATTGGAAATGTATAAACAGCACAG	180
Db	365	CGTTTGTCACTTTTCCAAATCCAGTACCTGTATATTGGAAATGTATAAACAGCACAG	424
QY	181	GCTTCTTCTGGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCAGGAGGAT	240
Db	425	GCTTCTTCTGGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCAGGAGGAT	484
QY	241	CTTAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCCCTTTTTCAGCAGGAT	300
Db	485	CTTAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCCCTTTTTCAGCAGGAT	544
QY	301	GGAAATGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGGTTCCAGAGCT	360
Db	545	GGAAATGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGGTTCCAGAGCT	604
QY	361	CGCTGTTCTATGGCTTTCAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTG	420
Db	605	CGCTGTTCTATGGCTTTCAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTG	664
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGAATGAA	480

Db	1265	ACAGATAACGTCTTCCACCTTGGATGCAGATTTT	1297
RESULT 14			
BD093078	Novel protein and its DNA.		
LOCUS	BD093078	4955 bp	DNA linear PAT 27-AUG-2002
DEFINITION	Novel protein and its DNA.		
ACCESSION	BD093078		
VERSION	BD093078.1 GI:22638666		
KEYWORDS	WO 0100799-A/2.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 4955)		
AUTHORS	Nakamura,Y., Arakawa,H. and Tanaka,H.		
TITLE	Novel protein and its DNA		
JOURNAL	Patent: WO 0100799-A 2 04-JAN-2001;		
TAKEDA CHEMICAL INDUSTRIES LTD,YUSUKE NAKAMURA,HIROFUMI ARAKAWA, HIROSHI TANAKA			
COMMENT	OS Homo sapiens (human)		
	PN	WO 0100799-A/2	
	PD	04-JAN-2001	
	PF	27-JUN-2000 WO 2000JP004189	
	PR	28-JUN-1999 JP 99P 181131,06-JUL-1999 JP 99P 192391 PR	
	PI	YUSUKE NAKAMURA,HIROFUMI ARAKAWA,HIROSHI TANAKA PC	
	C12N9/04,C12N15/53,C12N1/15,C12N1/19,C12N1/21,C12N5/10 PC		
	C12P21/02,A61K38/44,		
	PC A61K48/00,C07K16/40,G01N33/53,C12Q1/26		
	CC		
PH	Key	Location/Qualifiers.	
FEATURES	Location/Qualifiers		
source	1..4955		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match 99.8%; Score 1051.4; DB 6; Length 4955;			
Best Local Similarity 99.9%; Pred. No. 1.5e-254;			
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	1	ATGGGGGACCGGAAAGCGCGGCGGCTGGATCAGATGAGATCATCTTCA	60
Db	245	ATGGGGGACCGGAAAGCGCGGCGGCTGGATCAGATGAGATCATCTTCA	304
QY	61	GACACCAACGAAAGTCAAATAAAGTCAAATGAAGACCACTCCTAAGAAAGAGTTCTCGC	120
Db	305	GACACCAACGAAAGTCAAATAAAGTCAAATGAAGACCACTCCTAAGAAAGAGTTCTCGC	364
QY	121	CGGTTTGTCACTTTTCCAAATCCAGTACCTGTATATTGGAAATGTATAAACAGCACAG	180
Db	365	CGGTTTGTCACTTTTCCAAATCCAGTACCTGTATATTGGAAATGTATAAACAGCACAG	424
QY	181	GCTTCTCTTCTGGACAGCAGAGGTTGACTTATCAAAGGATCTCCCTCAGGAGGAT	240
Db	425	GCTTCTCTTCTGGACAGCAGAGGTTGACTTATCAAAGGATCTCCCTCAGGAGGAT	484
QY	241	CTTAAGCAGATGAGAAGTACTTCACTCTCACTTTAGCCCTTTTTCAGCAGGATGAT	300
Db	485	CTTAAGCAGATGAGAAGTACTTCACTCTCACTTTAGCCCTTTTTCAGCAGGATGAT	544
QY	301	GGAAATGTAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTCAGGTTCCAGAGCT	360
Db	545	GGAAATGTAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTCAGGTTCCAGAGCT	604
QY	361	CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG	420
Db	605	CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG	664
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA	480

665	CTGATAGACACTTACATCAGAGATCCCAGAAAAGGGAATTTTATTAAATGCAATTGAA	724
481	ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTGGCATGATGATAGCAGATAGAAA	540
725	ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTGGCATGATGATAGCAGATAGAAA	784
541	TCTACTTTTGGGAAAGAGTGGTGCCCTTGTCTGTGTAGAGGAGTTTCTTCTCAGGA	600
785	TCTACTTTTGGGAAAGAGTGGTGCCCTTGTCTGTGTAGAGGAGTTTCTTCTCAGGA	844
601	TCCTTTTGCTGTATATTCGTCTAAGAGAGAGTCTTATGCCCAGGACTCACTTTTTC	660
845	TCCTTTTGCTGTATATTCGTCTAAGAGAGAGTCTTATGCCCAGGACTCACTTTTTC	904
661	AATGAATCATCAGCAGAGATGAAGGACTTCACCTGTGACTTTGCTTGCCTGATGTTCAA	720
905	AATGAATCATCAGCAGAGATGAAGGACTTCACCTGTGACTTTGCTTGCCTGATGTTCAA	964
721	TACTTAGTAAATAAGCCCTTCAGAAAAAGGCTCAGGAGATCATTTGTGATGCTGTCAA	780
965	TACTTAGTAAATAAGCCCTTCAGAAAAAGGCTCAGGAGATCATTTGTGATGCTGTCAA	1024
781	ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGCCCTCATTTGGAATGAATTGCATT	840
1025	ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGCCCTCATTTGGAATGAATTGCATT	1084
841	TTGATGAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGGAATTTGGATTCTCA	900
1085	TTGATGAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGGAATTTGGATTCTCA	1144
901	AAGGTTTTTCAGGCAGAAAAATCCTTTTGATTTATGGAACAATTTCTTTAGAAGAAA	960
1145	AAGGTTTTTCAGGCAGAAAAATCCTTTTGATTTATGGAACAATTTCTTTAGAAGAAA	1204
961	ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTGCAAGTTATGCGAGAAC	1020
1205	ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTGCAAGTTATGCGAGAAC	1264
1021	ACAGATAACGCTTCACCTTGGATGACGATTTT	1053
1265	ACAGATAACGCTTCACCTTGGATGACGATTTT	1297
RESULT 15		
CUS	036063	
DEFINITION	Homo sapiens p53R2 mRNA for ribonucleotide reductase, complete cds.	
ACCESSION	AB036063	
VERSION	AB036063.1 GI:7229085	
KEYWORDS	p53R2; ribonucleotide reductase.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	Tanaka,H., Arakawa,H., Yamaguchi,T., Shiraishi,K., Fukuda,S., Matsui,K., Takei,Y. and Nakamura,Y.	
TITLE	A ribonucleotide reductase gene involved in a p53-dependent cell-cycle checkpoint for DNA damage	
JOURNAL	Nature 404 (6773), 42-49 (2000)	
MEDLINE	20179179	
PUBMED	10716435	
REFERENCE	2 (bases 1 to 4955)	
AUTHORS	Tanaka,H., Arakawa,H. and Nakamura,Y.	
TITLE	Direct Submission	
JOURNAL	Submitted (18-DEC-1999) Yusuke Nakamura, University of Tokyo, Institute of Medical Science, Human Genome Center, Laboratory of Molecular Medicine; 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan (E-mail:yusuke@ims.u-tokyo.ac.jp, Tel:+81-3-5449-5372, Fax:+81-3-5449-5433)	
FEATURES	Location/Qualifiers	
source	1. 4955 /organism="Homo sapiens"	

gene	/mol_type="mRNA" /db_xref="taxon:9606" /chromosome="8" /map="8q23.1" 1. .4955
CDS	/gene="p53R2" 245. .1300 /gene="p53R2" /codon_start=1 /product="ribonucleotide reductase" /protein_id="BAA92434.1" /db_xref="GI:7229086" /translation="MGDPERPEAAGLDQDERSSDNESEIKSNEPILRKSSRRVIVNIYVPIIDWKYKQAQSAFWTAEEVDLSKDLPHWNKLKADKDYFISHILAFFAASOGI VNEIYVRFESQVQVPEARCFYGFQILIEVHSEMYSLIDITYIRLKKREFLNFNAIE TMPYVKKADWALRWIADRKSTFGFVFAFAVEGVFFSGSFAAIFWLKKRGLPMPGIL FNEILISRDGLHCDPACILMFOYLVNKPSEERVREIIVDAKIEQGFLEALPVGLIG MNCILMKQVIEFVADRLVELGFSKVFQAEINPDPFMENTISLEKTNFPEKRVSEYQRF AVMAETIDNVFILDADF" 4933 /gene="p53R2" /note="22 a nucleotides"
polyA_site	
ORIGIN	
Query Match	99.8%; Score 1051.4; DB 9; Length 4955;
Best Local Similarity	99.9%; Pred. No. 1.5e-254;
Matches 1052;	Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	1 ATGGGGGACCCGGAAGCGCGGCGCGGCTGGATCAGATCAGATCAGATCATCTTCA 60
Db	245 ATGGGGGACCCGGAAGCGCGGCGCGGCTGGATCAGATCAGATCATCTTCA 304
QY	61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTCTCTCGC 120
Db	305 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTCTCTCGC 364
QY	121 CGGTTTGTTCATCTTTCCAAATCCAGTACCCTGTATTTTGGAAAAATGTATAAACAGGCACAG 180
Db	365 CGGTTTGTTCATCTTTCCAAATCCAGTACCCTGTATTTTGGAAAAATGTATAAACAGGCACAG 424
QY	181 GCTTTCCTTCTGGACACAGAGAGTTGACTTATCAAGAGATCTCCCTCACTCGAACACAG 240
Db	425 GCTTTCCTTCTGGACACAGAGAGTTCGACTTATCAAAAGGATCTCCCTCACTCGAACACAG 484
QY	241 CTTTAAAGCAGATGAGAAGTACTTTCATCTCTCATCATCTTAGCCCTTTTGTGAGCCAGCGTAT 300
Db	485 CTTTAAAGCAGATGAGAAGTACTTTCATCTCTCATCATCTTAGCCCTTTTGTGAGCCAGCGTAT 544
QY	301 GGAATTGTAAATGAAAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
Db	545 GGAATTGTAAATGAAAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 604
QY	361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAAGTTCATCTCAGAGATGTACAGTTTG 420
Db	605 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAAGTTCATCTCAGAGATGTACAGTTTG 664
QY	421 CTGTATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db	665 CTGTATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 724
QY	481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTGGCATGATGATAGCAGATAGAAA 540
Db	725 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTGGCATGATGATAGCAGATAGAAA 784
QY	541 TCTACTTTTGGGAAAGAGTGGTGCCCTTGTCTGTGTAGAGGAGTTTCTTCTCAGGA 600
Db	785 TCTACTTTTGGGAAAGAGTGGTGCCCTTGTCTGTGTAGAGGAGTTTCTTCTCAGGA 844
QY	601 TCCTTTTGCTGTATATTCGTCTAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Db	845 TCCTTTTGCTGTATATTCGTCTAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 904
QY	661 AATGAATCATCAGCAGAGATGAAGGACTTCACCTGTGACTTTGCTTGCCTGATGTTCCAA 720

/mol_type="mRNA"			
/db_xref="taxon:9606"			
/chromosome="8"			
/map="8q23.1"			
1..4955			
gene	/gene="p53R2"		
CDS	245..1300		
/gene="p53R2"			
/codon_start=1			
/product="ribonucleotide reductase"			
/protein_id="BAA92434.1"			
/db_xref="GI:7229086"			
/translation="MGDPERPEAAGLDODERSSDTNESEIKSNBPLLRKSSRRFVI FPIQYVDIWKYKQASFWTAEEVDLSKDLPHWNKLKADKVFISHILAFPAASDGI VNEILVERPSQEVQVPEARCFYGFQILLINHVSEMYSLIIDTYIRPKKREFLPNAIE TPMYKKADKALRIADKRSIFGERVAFVAVEGVFFSGSFAAIIFLKKRGLMPGLT FSNELISRDGLHCDLACLMFQYLNVKPSSEVRREIIVDAKIEQFLFEALPVLGIG NMCILMKQYIEFVADRLIVELGFSKVFQANPDPFMENTISLEKTNFFEKRVSEYORF AVMAETTDINVTLDADF"			
polyA_site	4933		
/gene="p53R2"			
/note="22 a nucleotides"			
ORIGIN			
Query Match 99.8%; Score 1051.4; DB 9; Length 4955;			
Best Local Similarity 99.9%; Pred. No. 1.5e-254;			
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Y	1	ATGGCGCAGCCCGGAAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA	60
b	245	ATGGCGCAGCCCGGAAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA	304
Y	61	GACACCAAGAAAGTGAATAAATCAATGAAGAGCCCTCTCTAAGAAAGAGTTCCTCGC	120
b	305	GACACCAAGAAAGTGAATAAATCAATGAAGAGCCCTCTCTAAGAAAGAGTTCCTCGC	364
Y	121	CGGTTTGTCTATCTTTCCAAATCCAGTACCTGATTTTGGAAAATGTATAAACAGGCACAG	180
b	365	CGGTTTGTCTATCTTTCCAAATCCAGTACCTGATTTTGGAAAATGTATAAACAGGCACAG	424
Y	181	GCTTCTCTTCTGGACAGCAGAGAGGTTGACTTATCAAAGGATCTCCCTCACTTGGAAACAAG	240
b	425	GCTTCTCTTCTGGACAGCAGAGAGGTTGACTTATCAAAGGATCTCCCTCACTTGGAAACAAG	484
Y	241	CTTAAAGCAGATGAGAAGTACTTCTATCTCTCACATCTTAGCCCTTTTTCAGCCAGCTGAT	300
b	485	CTTAAAGCAGATGAGAAGTACTTCTATCTCTCACATCTTAGCCCTTTTTCAGCCAGCTGAT	544
Y	301	GGAAATTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGGTTCCAGAGGCT	360
b	545	GGAAATTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGGTTCCAGAGGCT	604
Y	361	CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG	420
b	605	CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG	664
Y	421	CTGATAGACACTTATCATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA	480
b	665	CTGATAGACACTTATCATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA	724
Y	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTTGGCATGGATAGCAGATAGAAA	540
b	725	ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTTGGCATGGATAGCAGATAGAAA	784
Y	541	TCTACTTTTGGGAAAGAGTGGTGCCCTTTGCTGCTGTAGAGGAGTFTTCTTCTCAGGA	600
b	785	TCTACTTTTGGGAAAGAGTGGTGCCCTTTGCTGCTGTAGAGGAGTFTTCTTCTCAGGA	844
Y	601	TCCTTTGCTGCTATATTCCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC	660
b	845	TCCTTTGCTGCTATATTCCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC	904
Y	661	AATGAACACTCATCAGCAGATGAAGGACTTCACCTGTGACTTTTGCCTGATGTTCCAA	720

```
|||||
Db 905 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
QY 721 TACTTAGTAATAAGCCTTTCAAGAGAAAGGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 965 TACTTAGTAATAAGCCTTTCAAGAGAAAGGGTCAGGAGATCATTTGTTGATGCTGTCAAA 1024
QY 781 ATTGAGCAGGAGTTTTTAAACAGAAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTGCATT 840
Db 1025 ATTGAGCAGGAGTTTTTAAACAGAAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTGCATT 1084
QY 841 TTGATGAACACAGTACATTTGAGTTTGTAGCTGACAGATTACTTTGTGGAACTTGGATTCTCA 900
Db 1085 TTGATGAACACAGTACATTTGAGTTTGTAGCTGACAGATTACTTTGTGGAACTTGGATTCTCA 1144
QY 901 AAGGTTTTTTCAGGCAGAGAAATCCCTTTTGTATTTTATGGAAACATTTCTTTAGAGGAAA 960
Db 1145 AAGGTTTTTTCAGGCAGAGAAATCCCTTTTGTATTTTATGGAAACATTTCTTTAGAGGAAA 1204
QY 961 ACAATTTTCTTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
Db 1205 ACAATTTTCTTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1264
QY 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053
Db 1265 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1297
```

Search completed: October 30, 2005, 01:51:31
Job time : 3261 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2005, 21:38:30 ; Search time 440.5 Seconds
(without alignments)
14150.937 Million cell updates/sec

Title: US-10-698-228-12

Perfect score: 1053

Sequence: 1 atggcgaccggaaagcc.....tcaccttgatgcagatttt 1053

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1053	100.0	1053	4	Aaf32447 Human rib
2	1053	100.0	1081	4	Aaf32440 Human rib
3	1051.4	99.8	1053	4	Aaf32438 Human rib
4	1051.4	99.8	2596	3	Aaa12411 CDNA enco
5	1051.4	99.8	4955	4	Aaf32439 Human rib
6	1051.4	99.8	4955	13	Adr24210 Breast ca
7	1048.2	99.5	1601	4	Aah14924 Human cDN
8	650.6	61.8	706	13	Adq57092 Novel can
9	580	55.1	2482	4	Aas44917 Human con
10	580	55.1	2500	6	Abi65414 Lung canc
11	580	55.1	2500	6	Abi66517 Lung canc
12	580	55.1	2500	6	Abi65859 Lung canc
13	580	55.1	2500	8	Abx10335 DNA enco
14	580	55.1	2500	11	Adi32044 Human cDN
15	580	55.1	2500	12	Adn04443 Antipsori
16	580	55.1	2500	12	Adq09273 Human RRM
17	580	55.1	2500	13	Adn37637 Tumour-as
18	580	55.1	3393	12	Adn03788 Antipsori
19	580	55.1	3393	12	Ado19225 Human PRO
20	578.4	54.9	1989	3	Aac78111 Human can

21	578.4	54.9	2215	12	ADK70302
22	578.4	54.9	2216	10	ADJ56536
23	578.4	54.9	3294	4	Aah73225 Human cer
24	569	54.0	2641	4	Aas44745 Human ful
25	568.8	54.0	1170	2	AAV05641
26	562.6	53.4	1328	12	ADO57308
27	547.8	52.0	978	6	ABL90228
28	545.4	51.8	977	6	ABL90228 Human pol
29	527.2	50.1	608	4	AAH07707
30	514	48.8	1371	5	AAS79474
31	463	44.0	1289	4	ABL14627
32	442	42.0	1146	6	AAS62387
33	439.4	41.7	481	2	AAx51874
34	433.8	41.2	1218	8	ABT20895
35	429.8	40.8	14176	2	AAT84564
36	429.8	40.8	14176	4	AAF84949
37	429.8	40.8	14176	4	AAF84948
38	420.4	39.9	3945	4	ABL14626
39	404.8	38.4	963	13	ADS58378
40	398.6	37.9	1146	8	ABT19075
41	394.2	37.4	1173	13	ADT47814
42	393.6	37.4	1242	6	ABZ32250
43	388.2	36.9	1292	13	ADS49379
44	382	36.3	1200	10	ACC61238
45	382	36.3	1200	10	ADK63241

ALIGNMENTS

RESULT 1

AAF32447

ID AAF32447 standard; cDNA; 1053 BP.

XX

AC AAF32447;

XX

DT 18-APR-2001 (first entry)

XX

DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:12.

XX

KW Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.

XX

OS Homo sapiens.

XX

PN WO200100799-A1.

XX

PD 04-JAN-2001.

XX

PF 27-JUN-2000; 2000WO-JP004189.

XX

PR 28-JUN-1999; 99JP-00181131.

PR

06-JUL-1999; 99JP-00192391.

PR

21-JAN-2000; 2000JP-00017770.

XX

(TAKE) TAKEDA CHEM IND LTD.

PA

(NAKA/) NAKAMURA Y.

PA

Nakamura Y, Arakawa H, Tanaka H;

PI

WPI; 2001-112446/12.

DR

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX SQ Sequence 1053 BP; 314 A; 191 C; 246 G; 302 T; 0 U; 0 Other;

Query Match 100.0%; Score 1053; DB 4; Length 1053;
Best Local Similarity 100.0%; Pred. No. 9e-290;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGACCCGGAAGGCGGAGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 1 ATGGCGACCCGGAAGGCGGAGCGGCTGGATCAGGATGAGAGATCATCTTCA 60

QY 61 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTGC 120
DB 61 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTGC 120

QY 121 CGGTTGTTCATCTTCCAAATCAGTACCCCTGATATTTGGAAAAATGTATAACAGGCACAG 180
DB 121 CGGTTGTTCATCTTCCAAATCAGTACCCCTGATATTTGGAAAAATGTATAACAGGCACAG 180

QY 181 GCTTCCTTCTGCACAGCAGAGGTTGACTTATCAAAAGGATCTCCCTCACTGGAAACAG 240
DB 181 GCTTCCTTCTGCACAGCAGAGGTTGACTTATCAAAAGGATCTCCCTCACTGGAAACAG 240

QY 241 CTTAAGCAGATGAGAAGTACTTCTCATCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 300
DB 241 CTTAAGCAGATGAGAAGTACTTCTCATCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 300

QY 301 GGAATTGTAATGAATAATTTGTGAGCGCTTTAGTCAGGAGGTGCGAGGTTCAGAGGCT 360
DB 301 GGAATTGTAATGAATAATTTGTGAGCGCTTTAGTCAGGAGGTGCGAGGTTCAGAGGCT 360

QY 361 CGCTGTTCTATGGCTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGACAGTTG 420
DB 361 CGCTGTTCTATGGCTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGACAGTTG 420

QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAATTTTATTAATGCAATTGAA 480
DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAATTTTATTAATGCAATTGAA 480

QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCTTCGATGGATAGCAGATAGAAA 540
DB 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCTTCGATGGATAGCAGATAGAAA 540

QY 541 TCTACTTTTGGGAAGAGTGTGGCTTTGCTGTGTAGAGGAGTTTCTTCTCAGA 600
DB 541 TCTACTTTTGGGAAGAGTGTGGCTTTGCTGTGTAGAGGAGTTTCTTCTCAGA 600

QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGACTCACTTTTCC 660
DB 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGACTCACTTTTCC 660

QY 661 AATGAACATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGTATGTTCAA 720
DB 661 AATGAACATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGTATGTTCAA 720

QY 721 TACTTAGTAAATAGCCCTTCAGAAAGGGTTCAGGAGATCATTTGTGATGCTGCAAA 780
DB 721 TACTTAGTAAATAGCCCTTCAGAAAGGGTTCAGGAGATCATTTGTGATGCTGCAAA 780

QY 781 ATTGAGCAGAGGTTTTTAACAGAAGCTTGGCCCTCATTTGGAATGAATTGCATT 840
DB 781 ATTGAGCAGAGGTTTTTAACAGAAGCTTGGCCCTCATTTGGAATGAATTGCATT 840

QY 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAATCTGATTCTCA 900
DB 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAATCTGATTCTCA 900

QY 901 AAGGTTTTTCAGGCGAATAATCCTTTGATTTTATGGAATAATTTCTTTAGAGGAAA 960
DB 901 AAGGTTTTTCAGGCGAATAATCCTTTGATTTTATGGAATAATTTCTTTAGAGGAAA 960

QY 961 ACAAAATTTCTTGGAAAAAGGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAAGAAC 1020

DB 961 ACAAAATTTCTTGGAAAAAGGTTTCAGAGTATCAGCGTTTTCAGTATGCGAAGAAC 1020

QY 1021 ACAGATAACGCTCTTACACCTTGGATGAGATTTT 1053
DB 1021 ACAGATAACGCTCTTACACCTTGGATGAGATTTT 1053

RESULT 2
AAF32440
ID AAF32440 standard; cDNA; 1081 BP.
XX AAF32440;
XX AC
XX DT 18-APR-2001 (first entry)
XX DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:4.
XX KW Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
XX OS Homo sapiens.
XX PN WO200100799-A1.
XX PD 04-JAN-2001.
XX PF 27-JUN-2000; 2000WO-JP004189.
XX PR 28-JUN-1999; 99JP-00181131.
XX PR 06-JUL-1999; 99JP-00192391.
XX PR 21-JAN-2000; 2000JP-00017770.
XX (TAKE) TAKEDA CHEM IND LTD.
XX (NAKA/) NAKAMURA Y.
XX Nakamura Y, Arakawa H, Tanaka H;
XX WPI; 2001-112446/12.
XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
XX diagnosis, treatment and prevention of cancer.
XX Example 2; Page 91; 102pp; Japanese.
XX The present invention describes a human ribonucleotide reductase
XX designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of
XX the DNA repair mechanism and its activity is induced by p53. It can be
XX used for the treatment, prevention and diagnosis of a wide range of
XX cancers. The present sequence represents a human ribonucleotide reductase
XX related sequence which is used in an example from the present invention
XX SQ Sequence 1081 BP; 322 A; 199 C; 252 G; 308 T; 0 U; 0 Other;

Query Match 100.0%; Score 1053; DB 4; Length 1081;
Best Local Similarity 100.0%; Pred. No. 9.1e-290;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGCACCCGAAAGCGCGGAGCGGCTGGATCAGGATGAGATCATCTTCA 60
DB 20 ATGGGGCACCCGAAAGCGCGGAGCGGCTGGATCAGGATGAGATCATCTTCA 79

QY 61 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 80 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 139

QY 121 CGGTTTGTTCATCTTCCAAATCCAGTACCCCTGATATTTGGAAAAATGTATAACAGGCACAG 180
DB 140 CGGTTTGTTCATCTTCCAAATCCAGTACCCCTGATATTTGGAAAAATGTATAACAGGCACAG 199

QY 181 GCTTCCTTCTGGACAGCAGAAAGAGTTGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 240
DB 200 GCTTCCTTCTGGACAGCAGAAAGAGTTGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 259

QY 241 CTTAAGCAGATGAGAAGTACTTCACTCTTCACATCTTAGCCCTTTTTCAGCCAGTGAT 300

Db 260 CTTAAAGCAGATCAGAGTACTTCTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 319
 QY 301 GGAATTTGTAATAAGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTTCAGAGGCT 360
 Db 320 GGAATTTGTAATAAGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTTCAGAGGCT 379
 QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAAATGTTTCACTCAGAGATGTACAGTTTG 420
 Db 380 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAAATGTTTCACTCAGAGATGTACAGTTTG 439
 QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 480
 Db 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 499
 QY 481 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGCCCTTGGCATGGATAGCAGATAGAAA 540
 Db 500 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGCCCTTGGCATGGATAGCAGATAGAAA 559
 QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTTCCTTCTCAGGA 600
 Db 560 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTTCCTTCTCAGGA 619
 QY 601 TCTTTTGTCTGCTATATTCTGGCTAAAGAGAGAGTCTTATGCGAGACTCACTTTTTC 660
 Db 620 TCTTTTGTCTGCTATATTCTGGCTAAAGAGAGAGTCTTATGCGAGACTCACTTTTTC 679
 QY 661 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 720
 Db 680 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 739
 QY 721 TACTTAGTAATAAGCCCTTCAAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
 Db 740 TACTTAGTAATAAGCCCTTCAAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAA 799
 QY 781 ATTGACGACGAGTTTAAACAGAGCCCTTGCCAGTTGGCCCTCAATGGAATGATTCGATT 840
 Db 800 ATTGACGACGAGTTTAAACAGAGCCCTTGCCAGTTGGCCCTCAATGGAATGATTCGATT 859
 QY 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTGTGGAATCTGATTTCTCA 900
 Db 860 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTGTGGAATCTGATTTCTCA 919
 QY 901 AAGGTTTTCAGGCAAGAAATCCTTTTGTATTTATGGAACAAATTTCTTTAGAGGAAA 960
 Db 920 AAGGTTTTCAGGCAAGAAATCCTTTTGTATTTATGGAACAAATTTCTTTAGAGGAAA 979
 QY 961 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGTTATGGCAGAAACC 1020
 Db 980 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGTTATGGCAGAAACC 1039
 QY 1021 ACAGATAAGCTCTTACCTTGGATGCAGATTTT 1053
 Db 1040 ACAGATAAGCTCTTACCTTGGATGCAGATTTT 1072

RESULT 3

ID AAF32438 standard; cDNA; 1053 BP.

XX AAF32438;

AC AAF32438;

XX 18-APR-2001 (first entry)

DE Human ribonucleotide reductase TP53R2H nucleotide sequence SEQ ID NO:2.

XX Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.

OS Homo sapiens.

XX WO200100799-A1.

XX 04-JAN-2001.

XX 27-JUN-2000; 2000WO-JP004189.
 PF 28-JUN-1999; 99JP-00181131.
 PR 06-JUL-1999; 99JP-00192391.
 PR 21-JAN-2000; 2000JP-00017770.
 XX (TAKE) TAKEDA CHEM IND LTD.
 PA (NAKA/) NAKAMURA Y.
 XX Nakamura Y, Arakawa H, Tanaka H;
 PI WPI; 2001-112446/12.
 XX P-PSDB; AAB69050.
 DR
 DR
 XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
 PT diagnosis, treatment and prevention of cancer.
 XX Claim 5; Fig 1-3; 102pp; Japanese.
 XX The present sequence encodes a human ribonucleotide reductase designated
 CC TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of the DNA
 CC repair mechanism and its activity is induced by p53. It can be used for
 CC the treatment, prevention and diagnosis of a wide range of cancers
 XX
 SQ Sequence 1053 BP; 314 A; 192 C; 246 G; 301 T; 0 U; 0 Other;
 Query Match 99.8%; Score 1051.4; DB 4; Length 1053;
 Best Local Similarity 99.9%; Pred. No. 2.6e-289;
 Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGGGGACCCGGAAGGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
 Db 1 ATGGGGACCCGGAAGGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
 QY 61 GACACCAAGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
 Db 61 GACACCAAGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
 QY 121 CGGTTTGTCTCATCTTTCCAAATCCAGTACCTCGATATTTGGAAATGTATAACAGGCACAG 180
 Db 121 CGGTTTGTCTCATCTTTCCAAATCCAGTACCTCGATATTTGGAAATGTATAACAGGCACAG 180
 QY 181 GCTTCTCTCGACAGAGAGGTTGACTTATCAAAAGGATCTCCCTCACTGGAACAAG 240
 Db 181 GCTTCTCTCGACAGAGAGGTTGACTTATCAAAAGGATCTCCCTCACTGGAACAAG 240
 QY 241 CTTAAAGCAGATCAGAGTACTTTCATCTCTCAGATCTTAGCCTTTTTCGAGCCAGTGAT 300
 Db 241 CTTAAAGCAGATCAGAGTACTTTCATCTCTCAGATCTTAGCCTTTTTCGAGCCAGTGAT 300
 QY 301 GGAATTTGTAATAAGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGCT 360
 Db 301 GGAATTTGTAATAAGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGCT 360
 QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAAATGTTTCACTCAGAGATGTACAGTTTG 420
 Db 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAAATGTTTCACTCAGAGATGTACAGTTTG 420
 QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 480
 Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 480
 QY 481 ACCATGCCCTATGTTAAGAAAGAGAGTGGCCCTTGGCATGGATAGCAGATAGAAA 540
 Db 481 ACCATGCCCTATGTTAAGAAAGAGAGTGGCCCTTGGCATGGATAGCAGATAGAAA 540
 QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTTCCTTCTCAGGA 600
 Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTTCCTTCTCAGGA 600
 QY 601 TCTTTTGTCTGCTATATTCTGGCTAAAGAGAGGTTTATGCGAGACTCACTTTTTC 660

Db 601 TCTTTTGCTGCTATATTCCTGGCTAAAGAGAGAGGCTTTATGCCAGGACTCATCTTTTCC 660
Qy 661 AATGAATCATCAGCAGATCAAGGACTTCACTGTGACTTTGCTTGCCTGTGATGTTCCAA 720
Db 661 AATGAATCATCAGCAGATCAAGGACTTCACTGTGACTTTGCTTGCCTGTGATGTTCCAA 720
Qy 721 TACTTAGTAAATAGCCTTTCAGAGAAAGGGTCAGGAGATCAITTTGATGCTGTCAAA 780
Db 721 TACTTAGTAAATAGCCTTTCAGAGAAAGGGTCAGGAGATCAITTTGATGCTGTCAAA 780
Qy 781 ATTGACGAGGATTTTAAACAGAGCCTTGCCAGTTGGCTCATTTGGAAATGAATTGCATT 840
Db 781 ATTGACGAGGATTTTAAACAGAGCCTTGCCAGTTGGCTCATTTGGAAATGAATTGCATT 840
Qy 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTACTTTGTGAACTTCGATTCTCA 900
Db 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTACTTTGTGAACTTCGATTCTCA 900
Qy 901 AAGGTTTTTCAGCGAGAAATCCTTTTGTATTTTATGGAACATTTCTTTAGAAGGAAA 960
Db 901 AAGGTTTTTCAGCGAGAAATCCTTTTGTATTTTATGGAACATTTCTTTAGAAGGAAA 960
Qy 961 ACAAAATTTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
Db 961 ACAAAATTTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
Qy 1021 ACAGATAAGCTCTTCACCTTGATGACAGATT 1053
Db 1021 ACAGATAAGCTCTTCACCTTGATGACAGATT 1053

RESULT 4

AAA12411

ID AAA12411 standard; cDNA; 2596 BP.

XX AC AAA12411;

XX DT 25-JUL-2000 (first entry)

XX DE cDNA encoding a human RNA-associated protein.

XX KW Human; RNA-associated protein; cell proliferation; cancer; inflammation;
XX KW immune response; reproductive disorder; actinic keratosis;
XX KW atherosclerosis; arteriosclerosis; bursitis; cirrhosis; hepatitis;
XX KW mixed connective tissue disease; myelofibrosis; primary thrombocythemia;
XX KW paroxysmal nocturnal hemoglobinuria; polycythemia vera; psoriasis;
XX KW trauma; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT CDS 56..1111

XX FT /*tag= a

XX FT /product= "RNA-associated protein"

XX PN WO200015799-A2.

XX PD 23-MAR-2000.

XX PF 17-SEP-1999; 99WO-US021688.

XX PR 17-SEP-1998; 98US-00156039.

XX PR 22-SEP-1998; 98US-00158720.

XX PR 04-NOV-1998; 98US-00186815.

XX PR 08-APR-1999; 99US-0128660P.

XX PA (INCY-) INCYTE PHARM INC.

XX PI Tang YT, Corley NC, Guegler KJ, Gorgone GA, Patterson C;

XX PI Hillman JL, Baughn MR, Lal P, Azimzai Y, Yue H, Yang J;

XX DR WPI; 2000-271437/23.

XX DR P-PSDB; AAY84439.

XX New polypeptides and polynucleotides, useful for preventing and treating
PT a disorder associated with increased or decreased expression of RNA
PT associated proteins.

PS Claim 9; Page 120-121; 131pp; English.

XX The present sequence encodes a human RNA-associated protein. The
CC expression of RNA-associated proteins is closely associated with
CC reproductive tissues, nervous tissues, cell proliferation including
CC cancer, inflammation and immune responses, and so they may be used for
CC diagnosis, treatment or prevention of cell proliferative,
CC immune/inflammatory disorders, and reproductive disorders. Diseases and
CC disorders which may be treated include actinic keratosis,
CC atherosclerosis, arteriosclerosis, bursitis, cirrhosis, hepatitis, mixed
CC connective tissue disease, myelofibrosis, paroxysmal nocturnal
CC hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia
CC and cancers, and trauma

XX SQ Sequence 2596 BP; 783 A; 418 C; 510 G; 885 T; 0 U; 0 Other;

Query Match 99.8%; Score 1051.4; DB 3; Length 2596;

Best Local Similarity 99.9%; Pred. No. 3.9e-289;

Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGGCGACCCGAAAGCGCGGCGCGGCTGGATCAGATCAGAGATCATCTTCA 60

Db 56 ATGGGCGACCCGAAAGCGCGGCGCGGCTGGATCAGATCAGAGATCATCTTCA 115

Qy 61 GACACCAAGAAAGTAAAGTCAAAATGAAGACCACTCTCTAAGAAAGATTTCTCGC 120

Db 116 GACACCAAGAAAGTAAAGTCAAAATGAAGACCACTCTCTAAGAAAGATTTCTCGC 175

Qy 121 CGGTTTGTCTCTTCCAAATCCAGTACCTGTATATTTGGAAATGTATAACAGGCACAG 180

Db 176 CGGTTTGTCTCTTCCAAATCCAGTACCTGTATATTTGGAAATGTATAACAGGCACAG 235

Qy 181 GCTTCTCTCTGACAGCAGAGAGGTGTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240

Db 236 GCTTCTCTCTGACAGCAGAGAGGTGTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 295

Qy 241 CTTAAGCAGATGAGAAGTACTTCTCATCTTAGCCCTTTTTCGAGCCAGTGAT 300

Db 296 CTTAAGCAGATGAGAAGTACTTCTCATCTTAGCCCTTTTTCGAGCCAGTGAT 355

Qy 301 GGAATTGTAAATGAAATTTGTTGGAGCGCTTTAGTCAGGAGGTGCGAGTTCCAGAGGCT 360

Db 356 GGAATTGTAAATGAAATTTGTTGGAGCGCTTTAGTCAGGAGGTGCGAGTTCCAGAGGCT 415

Qy 361 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG 420

Db 416 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG 475

Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTGAA 480

Db 476 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTGAA 535

Qy 481 ACCATGCCCTATGTTTAAAGAAAAAGCAGATTGGGCCCTTCGATGGATAGCAGATAGAAA 540

Db 536 ACCATGCCCTATGTTTAAAGAAAAAGCAGATTGGGCCCTTCGATGGATAGCAGATAGAAA 595

Qy 541 TCTACTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600

Db 596 TCTACTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 655

Qy 601 TCTTTTGTCTGTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCCAGAGCTCACTTTTTC 660

Db 656 TCTTTTGTCTGTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCCAGAGCTCACTTTTTC 715

Qy 661 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGTGATGTTCCAA 720

Db 716 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGTGATGTTCCAA 775

QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAGCGGTGAGGAGATCATTTGTTGATGCTGTCAA 780
 Db TACTTAGTAAATAAGCCTTCAGAAAGAGCGGTGAGGAGATCATTTGTTGATGCTGTCAA 835
 QY 781 ATTGACGAGAGTTTAAACAGAGCCTTGCCAGTTGGCTCATTTGGAATGAATTCGATT 840
 Db ATTGACGAGAGTTTAAACAGAGCCTTGCCAGTTGGCTCATTTGGAATGAATTCGATT 895
 QY 841 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTAATCTTGTGGAACCTTGTCTCA 900
 Db TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTAATCTTGTGGAACCTTGTCTCA 955
 QY 901 AAGTTTTCAGCAGAAATCCTTTTGTATTTATGGAACATTTCTTTAGAGGAA 960
 Db AAGTTTTCAGCAGAAATCCTTTTGTATTTATGGAACATTTCTTTAGAGGAA 1015
 QY 961 ACAATTTCTTTCAGAAACAGTTTCAGATATCAGGTTTTCAGTTATGCGAGAAACC 1020
 Db ACAATTTCTTTCAGAAACAGTTTCAGATATCAGGTTTTCAGTTATGCGAGAAACC 1075
 QY 1021 ACAGATAAGCTCTTACCTTGGATGAGATTTT 1053
 Db ACAGATAAGCTCTTACCTTGGATGAGATTTT 1108

RESULT 5

AAF32439
 ID AAF32439 standard; cDNA; 4955 BP.

XX AAF32439;
 XX
 DT 18-APR-2001 (first entry)
 XX

DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:3.
 KW Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
 XX Homo sapiens.

XX WO200100799-A1.
 XX
 PD 04-JAN-2001.
 XX

XX 27-JUN-2000; 2000WO-JP004189.
 XX
 PR 28-JUN-1999; 99JP-00181131.
 PR 06-JUL-1999; 99JP-00192391.
 PR 21-JAN-2000; 2000JP-00017770.

XX (TAKE) TAKEDA CHEM IND LTD.
 PA (NAXA/) NAKAMURA Y.
 PA
 XX Nakamura Y, Arakawa H, Tanaka H;
 PI WPI; 2001-112446/12.

XX
 DR
 XX
 PT Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
 PT diagnosis, treatment and prevention of cancer.
 XX
 XX Example 2; Page 87-90; 102pp; Japanese.

XX The present invention describes a human ribonucleotide reductase
 CC designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of
 CC the DNA repair mechanism and its activity is induced by p53. It can be
 CC used for the treatment, prevention and diagnosis of a wide range of
 CC cancers. The present sequence represents a human ribonucleotide reductase
 CC related sequence which is used in an example from the present invention

XX Sequence 4955 BP; 1526 A; 803 C; 999 G; 1627 T; 0 U; 0 Other;
 Query Match 99.8%; Score 1051.4; DB 4; Length 4955;
 Best Local Similarity 99.9%; Pred. No. 5.2e-289;
 Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGACCCCGAAAGCGCGGAGCGGCGGTGGATCAGGATGAGAGATCATCTTCA 60
 Db ATGGGACCCCGAAAGCGCGGAGCGGCGGTGGATCAGGATGAGAGATCATCTTCA 304
 QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACCTCTTAAGAAAGAGTTCTCGC 120
 Db GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACCTCTTAAGAAAGAGTTCTCGC 364
 QY 121 CGGTTTGTCTATCTTCCAAATCCAGTACCTCTGATATTTGGAAATGTATAAAGCAGCAG 180
 Db CGGTTTGTCTATCTTCCAAATCCAGTACCTCTGATATTTGGAAATGTATAAAGCAGCAG 424
 QY 181 GCTTCTCTTCGACACGACGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
 Db GCTTCTCTTCGACACGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 484
 QY 241 CTTAAAGCAGATCAGAAAGTACTTCACTCTCACTCTTAGCCTTTTTCAGCCAGTGTAT 300
 Db CTTAAAGCAGATCAGAAAGTACTTCACTCTCACTCTTAGCCTTTTTCAGCCAGTGTAT 544
 QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGTCAGGTTCCAGAGCT 360
 Db GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGTCAGGTTCCAGAGCT 604
 QY 361 CGCTGTTCTATGGCTTCAAAATCTCATCGAGATGTTCTCACTCAGAGATGTACAGTTTG 420
 Db CGCTGTTCTATGGCTTCAAAATCTCATCGAGATGTTCTCACTCAGAGATGTACAGTTTG 664
 QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATGAA 480
 Db CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATGAA 724
 QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTTCGATGGATAGCAGATAGAAAA 540
 Db ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTTCGATGGATAGCAGATAGAAAA 784
 QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGATTTTCTTCTCAGGA 600
 Db TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGATTTTCTTCTCAGGA 844
 QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGCTTATGCCAGAGCTACCTTTTTC 660
 Db TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGCTTATGCCAGAGCTACCTTTTTC 904
 QY 661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
 Db AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
 QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTGAGGAGATCATTTGTTGATGCTGTCAA 780
 Db TACTTAGTAAATAAGCCTTCAGAAAGAGGTGAGGAGATCATTTGTTGATGCTGTCAA 1024
 QY 781 ATTGACGAGAGTTTAAACAGAGCCTTGCCAGTTGGCTCATTTGGAATGAATTCGATT 840
 Db ATTGACGAGAGTTTAAACAGAGCCTTGCCAGTTGGCTCATTTGGAATGAATTCGATT 1084
 QY 841 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTAATCTTGTGGAACCTTGTCTCA 900
 Db TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTAATCTTGTGGAACCTTGTCTCA 1144
 QY 901 AAGTTTTCAGCAGAAATCCTTTTGTATTTATGGAACATTTCTTTAGAGGAA 960
 Db AAGTTTTCAGCAGAAATCCTTTTGTATTTATGGAACATTTCTTTAGAGGAA 1204
 QY 961 ACAATTTCTTTCAGAAACAGTTTCAGATATCAGGTTTTCAGTTATGCGAGAAACC 1020
 Db ACAATTTCTTTCAGAAACAGTTTCAGATATCAGGTTTTCAGTTATGCGAGAAACC 1264
 QY 1021 ACAGATAAGCTCTTACCTTGGATGAGATTTT 1053
 Db ACAGATAAGCTCTTACCTTGGATGAGATTTT 1297

XX (GENE-) GENE LOGIC INC.
PA (PFIZ) PFIZER PROD INC.
XX
PI Diggins JC, Porter M, Wei T;
XX WPI; 2004-561890/54.
XX
XX New isolated nucleic acid molecule, useful for drug screening and
PT toxicity assays or for assessing the impact, including toxicity, of a
PT compound, pharmaceutical agent or environmental pollutant on a cell or
PT living organism.
XX
PS Claim 1; SEQ ID NO 8394; 41pp; English.
XX
XX This invention is related to a novel isolated canine nucleic acid
CC sequences and the construction of canine microarrays containing a
CC significant portion of the canine genome. The isolated canine nucleic
CC acid sequences of the invention may be useful for drug screening and
CC toxicity assays. The invention is therefore useful for assessing the
CC impact, including toxicity, of a compound, pharmaceutical agent or
CC environmental pollutant on a cell or living organism. The methods are
CC useful for detecting genes that are up- or down-regulated in canines in a
CC disease state. The sequences are useful as diagnostic agents or markers
CC to detect a cellular response in a sample individually or as part of a
CC gene expression profile. It is also useful as a target for agents that
CC modulate gene expression or activity. The database is useful for
CC producing electronic Northern blots that allow the user to determine the cell
CC type or tissue in which a given gene is expressed and to allow
CC determination of the abundance or expression level of a given gene in a
CC particular tissue or cell. The methods are useful for determining the
CC similarity of a toxic response to one or more individual compounds. The
CC methods are useful for predicting at least one toxic response or the
CC likelihood that a compound or test agent will induce various specific
CC pathologies such as those of the liver (liver necrosis, fatty liver
CC disease, protein adduct formation or hepatitis), those of the kidney,
CC heart, brain or testes, or other pathologies associated with at least one
CC of the toxins. The methods are also useful for predicting or elucidating
CC the potential cellular pathways influenced, induced or modulated by the
CC compound or test agent due to the similarity of the expression profile
CC compared to the profile induced by a known toxin. The present sequence is
CC that of a canine DNA sequence which was claimed for use during the
CC production of a canine microarray of the invention.
XX
SQ Sequence 706 BP; 208 A; 113 C; 158 G; 227 T; 0 U; 0 Other;

Query Match 61.8%; Score 650.6; DB 13; Length 706;
Best Local Similarity 95.2%; Pred. No. 5.5e-175;
Matches 671; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 349 GTTCCAGAGGCTCGCTGTTTCTATGGCTTTTCAAAATCTCATCGAAGATGTTCACTCAGAG 408
DB 1 GTTCCAGAGGCTCGCTGTTTCTATGGCTTTTCAAAATCTCATCGAAGATGTTCACTCAGAG 60
QY 409 ATGTACAGTTTCTGATACACTTACATCAGAGATCCCAAGAAAGGAATTTTATTT 468
DB 61 ATGTACAGTTTCTGATACACTTACATCAGAGATCCCAAGAAAGGAATTTTATTT 120
QY 469 AATGCAATTTGAAACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCCCTTGCATGGATA 528
DB 121 AATGCAATTTGAAACCATGCCATATGTTAAGAAAAAGCAGATTGGGCCCTTGCATGGATA 180
QY 529 GCAGATAGAAAATCTACTTTTGGGAAAGAGTGGTGGCCCTTTGCTGCTGTAGAAGAGTT 588
DB 181 GAAGATAGAAAATCTACTTTTGGGAAAGAGTGGTGGCCCTTTGCTGCTGTAGAAGAGTT 240
QY 589 TTTCTTCTCAGGATCTTTTCTGCTATATTTCTGGCTAAAGACAGAGGTTTATGCGAGA 648
DB 241 TTTCTTCTCAGGATCTTTTCTGCTATATTTCTGGCTAAAGACAGAGGTTTATGCGCTGA 300
QY 649 CTCACCTTTTCCAAATGAACCTCATCAGCAGAGATGAAGGACTTTCACGTGACTTTGCTTCC 708
DB 301 CTCACCTTTTCCAAATGAACCTCATCAGCAGAGATGAAGGACTTTCACGTGACTTTGCTTCC 360

QY 709 CTGATCTTCCAATACTTAGTAAATAGCCCTTCAGAAAGAGGCTCAGGGAGATCATTTGTT 768
DB 361 CTGATCTTCCAATACTTAGTAAATAGCCCTTCAGAAAGAGGCTCAGGGAGATCATTTGTT 420
QY 769 GATGCTGTCAAAATTTGAGCAGGAGTGTAAACACAGAACCCCTTGCAGTTGSCCTCATTTGGA 828
DB 421 AATGCTGTTGAAATTTGAGCAGGAGTGTAAACCGAAGCCCTTGCAGTTGSCCTCATTTGGA 480
QY 829 ATGAATTCATTTTTCATGAAACAGATCATTTGTTAGCTGACAGATTTACTTTGGGAA 888
DB 481 ATGAATTTGTTTTCATGAAACAGATCATTTTACTTTGTTAGCTGACAGATTTACTTTGGGAA 540
QY 889 CTTGGATCTCTCAAGGTTTTCAGGCAGAGAAATTCCTTTTATGATTTATGAAAAACATTTCT 948
DB 541 CTTGGATCTCTCAAGGTTTTCAGGCAGAGAAATTCCTTTTATGATTTATGAAAAACATTTCT 600
QY 949 TTAGAAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGAGTT 1008
DB 601 TTAGAAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGCGTT 660
QY 1009 ATGGCAGAAACCCACAGATTAACGTTCTTCACCTTCGATGTCAGATTTT 1053
DB 661 ATGGCAGAAACCCACAGATTAACGTTCTTCACCTTCGATGTCAGATTTT 705

RESULT 9
AAS44917/c
ID AAS44917 standard; DNA; 2482 BP.
XX
AC AAS44917;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human contig polynucleotide sequence #170.
XX
KW Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
KW nervous system disorder; inflammatory disorder; cell differentiation; ds;
KW angiogenesis; stem cell growth factor; activin; inhibitor; cartilage; burn;
KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;
KW cystostatic; antirheumatic; antiarthritic; vulnery; antiinflammatory;
KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
KW neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;
KW immunostimulant; analgesic; gene therapy.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200164834-A2.
XX
PD 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US004926.
PF
XX 28-FEB-2000; 2000US-00515126.
XX 18-MAY-2000; 2000US-00577409.
PR 17-JUN-2000; 2000US-00597707.
PR 14-JUL-2000; 2000US-00616807.
PR 19-SEP-2000; 2000US-00664641.
XX
PA (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;
PI Drmanac R;
XX
XX WPI; 2001-589862/66.
DR P-PSDB; AAU28017.
XX
XX Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis, treatment of cancer,
PT

PT neurological, inflammatory disorders and for use in arrays for detection.
XX
PS Claim 1; SEQ ID NO 514; 153bp; English.
XX

CC Sequences AAS44576-AAS44919 represent full-length polynucleotides and
CC contig polynucleotides encoding polypeptides of the invention. The DNA
CC and protein sequences are useful for the treatment, diagnosis and
CC prevention of various types of disorder in a mammalian subject such as a
CC human, dog, monkey, mouse, hamster or rat. The disorders include cancers
CC such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such
CC as multiple sclerosis, connective tissue disease, rheumatoid arthritis,
CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's
CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
CC bowel disease. The sequences exhibit activity relating to angiogenesis,
CC cell proliferation, cell differentiation, stem cell growth factor,
CC activin or inhibin. Therefore, they can be used to manipulate stem cells
CC in culture to give rise to neuroepithelial cells that can be used to
CC augment or replace cells damaged by illness, accidental damage or genetic
CC disorders. The sequences may also be used for regeneration of bone,
CC cartilage, tendons and ligaments and in tissue repair and burn healing.
CC Note: Some sequences for this patent did not form part of the printed
CC specification, but were obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 2482 BP; 707 A; 598 C; 522 G; 655 T; 0 U; 0 Other;

Query Match 55.1%; Score 580; DB 4; Length 2482;
Best Local Similarity 75.1%; Pred. No. 1.4e-154;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 89 ATGAAGAGGACCTCTTAAGAAAGATCTCGCGCGTTTGTTCATCTTTCCAAATCCAGTACC 148
DB 2086 AGGATGAGCGCTGTGAGAGAAACCCCGCGCTTTGTTCATCTTTCCCAATCCAGTACC 2027
QY 149 CTGATATTTGGAAATGTATAACAGGACAGGCTTCTCTGGACAGAGAGGTTG 208
DB 2026 ATGATATCTGGCAGATGTATAGAGGAGAGGCTTCTTTGGACCGCGAGAGGTTG 1967
QY 209 ACTTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAGTACTTCATCT 268
DB 1966 ACCTCTCCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGATATTTATAT 1907
QY 269 CTCACATCTTAGCCTTTTGGCCAGGATGATGGAATGTAAATGAAATTTGGTGGAGC 328
DB 1906 CCCATGCTTGGCTTTCTTTGGCAGCAAGCGATGGCATAGTAAATGAAACTTGGTGGAGC 1847
QY 329 GCTTTAGTCAGGAGGTGAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTCAAAATCTCA 388
DB 1846 GATTTAGCCAAAGATTCAGATTACAGAACCCCGCTGTTTCTATGGCTTCCAAATGGCCA 1787
QY 389 TCAGAAATGTTCACTCAGAGATGTACAGTTTGTGTATAGACACTTACATCAGAGATCCCA 448
DB 1786 TGGAAACATACATCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 1727
QY 449 AGAAAGGGAATTTTATTAATGCAATTTGAAACCATGCCCTATGTTTAAAGAAAAAGCAG 508
DB 1726 AAGAAAGGGAATTTCTCTTCAATGCCATTTGAAACGATGCTTGTGTCAAGAGAGGAGC 1667
QY 509 ATTGGCCCTTGGATGGATGAGATGAGAAATCTACTTTTGGGGAAGAGTGGTGGCCT 568
DB 1666 ACTGGCCCTTGGCTGGATTTGGGGAACAAAGAGGCTACCTATGGTGAACGCTTGTAGCCT 1607
QY 569 TTGCTGCTGTAAGAGGATTTTCTTCTCAGGATCTTTTGTCTGCTATATTCTGCTAAAGA 628
DB 1606 TTGCTGCTGTAAGAGGATTTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1547
QY 629 AGAGAGGCTTATGCGAGGACTCATTCTTTTCAATGAACCTATCAGCAGAGATGAGGAC 688
DB 1546 AACGAGGACTGATGCTGCGCCTCACATTTTCTAAATGAACCTTATAGCAGAGATGAGGTT 1487

QY 689 TTCACTGTGACTTTGCTTGGCTGATGTTTCCAACTACTAGTAAATAAGCCTTTCAGAAAGAA 748
DB 1486 TACACTGTGATTTTGTCTTGGCTGATGTTTCAACACCTGTTACACAAACATCGGAGGAGA 1427
QY 749 GGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTTGAGCAGGAGTTTTTAACAGAGCCT 808
DB 1426 GAGTAAGAGAAATAATTAATCAATGCTGTTCCGATAGAACAGGAGTTCCTCACTGAGGCCT 1367
QY 809 TGCAGTGGCCTCATTTGGAATGAATTTGATGAAACAGTACATTTGATTTGTTAG 868
DB 1366 TGCCTGTGAAGACTCATTTGGGATGAATTTGCACTCTAATGAAGCAATACATTTGATTTGTGG 1307
QY 869 CTGACAGATTAATTTGCTGAACTTCTGATTTCTCAAAAGTTTTTTCAGCAGAAATCTTTTG 928
DB 1306 CAGACAGCTTATGCTGGAATCTGGTTTTTACCAAGTTTTTTCAGAGTAGAGAACCATTTG 1247
QY 929 ATTTATGGAACATTTCTTTTGAAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
DB 1246 ACTTTATGGAGATATTTTCACTGGAAGAAAGACTTAACCTCTTTTGAAGAGAGTAGGCG 1187
QY 989 AGTATCAGCGTTTTCAGTTATGCGCAAAACCAAGATGAGTAAAGTCTTCACTTTGGATGCG 1048
DB 1186 AGTATCAGAGATGGGAGTGTGTCAAAGTCCACAGAGAAATTTCTTTTACCTTTGGATGCTG 1127
QY 1049 ATTT 1052
DB 1126 ACTT 1123

RESULT 10
ABL65414

ID ABL65414 standard; DNA; 2500 BP.

XX ABL65414;

XX 15-MAY-2002 (first entry)

XX Lung cancer related gene sequence SEQ ID NO:3751.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
XX cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
XX gene; ds.

XX Homo sapiens.

XX WO2001194629-A2.

XX 13-DEC-2001.

XX 30-MAY-2001; 2001WO-US010838.

XX 05-JUN-2000; 2000US-0209473P.

XX 05-JUN-2000; 2000US-0209531P.

XX 18-SEP-2000; 2000US-0233133P.

XX 18-SEP-2000; 2000US-0233617P.

XX 20-SEP-2000; 2000US-0234009P.

XX 20-SEP-2000; 2000US-0234034P.

XX 20-SEP-2000; 2000US-0234052P.

XX 22-SEP-2000; 2000US-0234509P.

XX 22-SEP-2000; 2000US-0234567P.

XX 25-SEP-2000; 2000US-0234923P.

XX 25-SEP-2000; 2000US-0234924P.

XX 25-SEP-2000; 2000US-0235077P.

XX 25-SEP-2000; 2000US-0235082P.

XX 25-SEP-2000; 2000US-0235134P.

XX 25-SEP-2000; 2000US-0235280P.

XX 26-SEP-2000; 2000US-0235637P.

XX 26-SEP-2000; 2000US-0235638P.

XX 27-SEP-2000; 2000US-0235711P.

XX 27-SEP-2000; 2000US-0235720P.

XX 27-SEP-2000; 2000US-0235840P.

XX 27-SEP-2000; 2000US-0235863P.

PR 28-SEP-2000; 2000US-0236028P.
 PR 28-SEP-2000; 2000US-0236032P.
 PR 28-SEP-2000; 2000US-0236033P.
 PR 28-SEP-2000; 2000US-0236034P.
 PR 28-SEP-2000; 2000US-0236109P.
 PR 28-SEP-2000; 2000US-0236111P.
 PR 29-SEP-2000; 2000US-0236842P.
 PR 29-SEP-2000; 2000US-0236891P.
 PR 02-OCT-2000; 2000US-0237172P.
 PR 02-OCT-2000; 2000US-0237173P.
 PR 02-OCT-2000; 2000US-0237278P.
 PR 02-OCT-2000; 2000US-0237294P.
 PR 02-OCT-2000; 2000US-0237295P.
 PR 02-OCT-2000; 2000US-0237316P.
 PR 03-OCT-2000; 2000US-0237425P.
 PR 03-OCT-2000; 2000US-0237598P.
 PR 03-OCT-2000; 2000US-0237604P.
 PR 03-OCT-2000; 2000US-0237606P.
 PR 03-OCT-2000; 2000US-0237608P.
 PR 01-NOV-2000; 2000US-0244867P.
 PR 01-NOV-2000; 2000US-0245084P.
 XX
 PA (AVAL-) AVALON PHARM.
 XX
 XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 PI
 XX WPI; 2002-188264/24.
 XX
 XX Screening for anti-neoplastic agent involves exposing cells to a chemical
 PT agent to be tested for anti-neoplastic activity, and determining a change
 PT in expression of a gene of a signature gene set.
 XX
 XX Claim 1; SEQ ID NO 3751; 44pp; English.
 PS
 XX The present invention describes a method (M1) for screening for an anti-
 CC neoplastic agent. The method involves exposing cells to a chemical agent
 CC to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL6164
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening an
 CC anti-neoplastic agent, and can be used for producing a product which is
 CC the data collected with respect to the anti-neoplastic agent as a result
 CC of M1, and the data is sufficient to convey the chemical structure and/or
 CC properties of the agent. M1 can be used in the treatment of cancer such
 CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
 CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
 CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
 CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms
 CC tumour
 XX
 SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;
 Query Match 55.1%; Score 580; DB 6; Length 2500;
 Best Local Similarity 75.1%; Pred. No. 1.4e-154;
 Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
 QY 89 ATGAGAGGACCTCTCTAAGAAAGAGTTCTCGCGCGGTTTGTCTATCTTCCAAATCCAGTACC 148
 DB 397 AGGATGAGCGCGTGTGAGAGAAACCCCGCGCGTTTGTCTATCTTCCCATCGAGTACC 456
 QY 149 CTGATATTGGAAAATGTATAACAGGCACAGGCTTCTCTTCTGGACAGAGAAGAGTTG 208
 DB 457 ATGATATCTGGCAGATGTATAGAGGCGAGAGGCTTCTCTTGGACCGCGAGGAGTTG 516
 QY 209 ACTTATCAAAGGATCTCCCTGACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTCATCT 268
 DB 517 ACCTCTCCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
 QY 269 CTCACATCTTAGCCCTTTTTCGAGCGAGTGATGGAATCTAAATTTGGTGGAGC 328

Db 577 CCCATGTTCTGGCTTTCTTTTGCAGCAAGCGATGGCATAGTAATAAGAAAACCTTGGTGAGC 636
 QY 329 GCTTTAGTCAGAGGTGCGAGGTTCCAGAGGCTCGCTGTTCTTATGCTTCAATTTCTCA 388
 Db 637 GATTTAGCCCAAGAAGTTTCAGATTACAGAAAGCCCGCTGTTCTATGCTTCCAAATTTGCCA 696
 QY 389 TCGAGAATGTTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCA 448
 Db 697 TGGAAAAACATACATCTGGAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 756
 QY 449 AGAAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCCCTATGTTAAGAAAAGAGCAG 508
 Db 757 AAGAAAAGGGAATTTCTCTTCAATGCCATTGAAACGATGCCCTTGTGTCAAGAAAGGAGCAG 816
 QY 509 ATTGGCCCTTGGATGGATAGCAGATAGAAAATCTACTTTTGGGGAAAGAGTGGTGCCT 568
 Db 817 ACTGGGCCCTTGGCGTGGATTGGGGACAAAGAGGCTACCTATGGTGAACGTTGTAGCCT 876
 QY 569 TTGCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTCTGCTGCTATATTCTGGCTAAAGA 628
 Db 877 TTGCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTCTGCTGCTATATTCTGGCTCAAGA 936
 QY 629 AGAGAGGTCTTATGCCAGGACTCACATTTTCCAACTCACTTAGTAATAAGCCCTTCAGAGAAA 688
 Db 937 AACGAGGACTGATGCCCTGCCCTCACATTTTCTTAATGAACTTATTAGCAGAGATGAGGTT 996
 QY 689 TTCACCTGTGACTTTGCTTGCCTGATGTTCCAACTACTTAGTAATAAGCCCTTCAGAGAAA 748
 Db 997 TACACTGTGATTTTGTCTGCTGATGTTCCAAACACCTGGGTACACAAACCAATCGAGGAGA 1056
 QY 749 GGGTCAGGAGAGATCATTTGTTGATGCTGTCAAATTCAGCAGGAGGTTTAAACAGAGCCCT 808
 Db 1057 GAGTAAGAGAAATATATCAATGCTGTTCCGATAGAACAGGAGTTCTCTCCTGAGGCT 1116
 QY 809 TGCCAGTTGGCCCTCATTTGGAATGAATTCGATTTTGAATAACAGTAGTACATTGAGTTGTAG 868
 Db 1117 TGCTGTGAAGCTCATTTGGATGAATTCGACTTAATGAAGCAATACATTGAGTTGTGG 1176
 QY 869 CTGACAGATTAATTTGGAACCTTGAATCTCAAAGGTTTTCAGGAGAGAAAATCCTTTTG 928
 Db 1177 CAGACAGACTTATGCTGGAACCTGGGTTTTCAGGAGGTTTTCAGAGTAGAGAACCCCATTTG 1236
 QY 929 ATTTTATGAAAACATTTCTTTAGAGGAGAAAACAAATTTCTTTTGAGAAAACGAGTTTTCAG 988
 Db 1237 ACTTATGAGAGATATTTTCTGGAAGGAAAGTAACTTCTTTTGAGAGAGATGAGCG 1296
 QY 989 AGTATCAGCGTTTTCAGATTATGGCAGAAACCAACAGATAACGCTCTTCACCTTGGATCAG 1048
 Db 1297 AGTATCAGAGGATGGGAGTGATGTCAGTCAAGTCCAACAGAGAAATCTTTTACCTTGGATCTG 1356
 QY 1049 ATTT 1052
 Db 1357 ACTT 1360

RESULT 11
 ABL66517
 ID ABL66517 standard; DNA; 2500 BP.
 XX
 AC ABL66517;
 XX
 DT 15-MAY-2002 (first entry)
 XX
 DE Lung cancer related gene sequence SEQ ID NO:4854.
 XX
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilms's tumour; adenocarcinoma;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200194629-A2.

XX PD
XX PF
XX PF
XX PF
13-DEC-2001.
30-MAY-2001; 2001WO-US010838.
05-JUN-2000; 2000US-0209473P.
05-JUN-2000; 2000US-0209531P.
18-SEP-2000; 2000US-0231333P.
18-SEP-2000; 2000US-0233617P.
20-SEP-2000; 2000US-0234009P.
20-SEP-2000; 2000US-0234034P.
20-SEP-2000; 2000US-0234052P.
22-SEP-2000; 2000US-0234509P.
22-SEP-2000; 2000US-0234567P.
25-SEP-2000; 2000US-0234923P.
25-SEP-2000; 2000US-0234924P.
25-SEP-2000; 2000US-0235077P.
25-SEP-2000; 2000US-0235082P.
25-SEP-2000; 2000US-0235134P.
25-SEP-2000; 2000US-0235280P.
26-SEP-2000; 2000US-0235637P.
26-SEP-2000; 2000US-0235638P.
27-SEP-2000; 2000US-0235711P.
27-SEP-2000; 2000US-0235720P.
27-SEP-2000; 2000US-0235840P.
27-SEP-2000; 2000US-0235863P.
28-SEP-2000; 2000US-0236028P.
28-SEP-2000; 2000US-0236032P.
28-SEP-2000; 2000US-0236033P.
28-SEP-2000; 2000US-0236034P.
28-SEP-2000; 2000US-0236109P.
28-SEP-2000; 2000US-0236111P.
29-SEP-2000; 2000US-0236842P.
29-SEP-2000; 2000US-0236891P.
02-OCT-2000; 2000US-0237172P.
02-OCT-2000; 2000US-0237173P.
02-OCT-2000; 2000US-0237278P.
02-OCT-2000; 2000US-0237294P.
02-OCT-2000; 2000US-0237295P.
02-OCT-2000; 2000US-0237316P.
03-OCT-2000; 2000US-0237425P.
03-OCT-2000; 2000US-0237598P.
03-OCT-2000; 2000US-0237604P.
03-OCT-2000; 2000US-0237606P.
03-OCT-2000; 2000US-0237608P.
01-NOV-2000; 2000US-0244867P.
01-NOV-2000; 2000US-0245084P.
(AVAL-) AVALON PHARM.
Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
Soppet DR, Weaver Z;
WPI; 2002-188264/24.
Screening for anti-neoplastic agent involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, and determining a change in expression of a gene of a signature gene set.
Claim 1; SEQ ID NO 4854; 44pp; English.
The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in ABL61664 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such

CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, squamous, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms
CC tumour
XX
SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;
Query Match 55.1%; Score 580; DB 6; Length 2500;
Best Local Similarity 75.1%; Pred. No. 1.4e-154;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
QY 89 ATGAAGAGCCACTCTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTCCCAATCCAGTACC 148
DB 397 AGGATGAGCGCTGCTGAGAGAAAACCCCGCGCTTGTTCATCTTCCCAATCCAGTACC 456
QY 149 CTGATATTTGGAAAAATGTATAAACAGGCACAGCTTCTCTTCTGGACAGCAAGAGTTG 208
DB 457 ATGATATCTGGCAGATGTATAAGAAAGCAGAGGCTTCTCTTGGACCGCGAGGTTG 516
QY 209 ACTTATCAAGAGATCTCCCTCAGTGAACAAAGCTTAAAGCAGATGAGAGTACTTCTCT 268
DB 517 ACCTCTCCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGATATTTTATAT 576
QY 269 CTCACATCTTAGCCTTTTGGCAGCCAGTGTATGGAATGTAAATGAAAAATTTGGTGAGC 328
DB 577 CCATGTTCTGGCTTCTTGGCAGCAAGCATGGCATAGTAATGAAAAATTTGGTGAGC 636
QY 329 GCTTTAGTCAGGAGTGCGAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTCCAAATCTCA 388
DB 637 GATTTAGCCAAGAAGTTCAGATTACAGAACCCGCTGTTTCTATGGCTTCCAAATGCCA 696
QY 389 TCGAAGATGTTCACTCAGAGATGTACAGTTGCTGATAGACACTTATCAGAGATCCCA 448
DB 697 TGGAAAAACATACATCTGAAATGTATAGTCTTCTTATGACACTTACATACAAAGATCCA 756
QY 449 AGAAAGGGAATTTTATTTAATGCAATGAACCATGCCCTATGTTTAAAGAAAAAGCAG 508
DB 757 AAGAAAGGGAATTTCTCTCAATGCCATTGAAACGATGCCCTTGTGTCAAGAAAGGACG 816
QY 509 ATTGGCCCTTGGCATGGATAGACAGATAGAAAAATCTACTTTTGGGAAAGAGTGTGGCCT 568
DB 817 ACTGGCCCTTGGCTGATTTGGGACAAAGAGGCTACCTATGTTGACGCTGTGTAGCCT 876
QY 569 TTGCTGCTGTAGAAAGAGTTTCTCTCAGAGATCTTTTGTGCTATATCTGGCTAAAGA 628
DB 877 TTGCTGAGTGAAGGCAATTTCTTTCCGGTTCTTTTGGCTGATATCTGGCTCAAGA 936
QY 629 AGAGAGGTCTTATGCCAGGACTCACTTTTCCCAATGAACTCATCAGCAGAGATGAAGAC 688
DB 937 AAGCAGGACTGATGCTCGCTGCTCACATTTTCTAATGAACCTTATAGCAGAGATGAGGTT 996
QY 689 TTCACCTGTGACTTTGCTTGGCTGATGTTCCAATCTTAGTAATAAGCTTTCAGAAAGAA 748
DB 997 TACACTGTGATTTGCTTGGCTGATGTTCAACACCTGTTACACAAACCATCGAGAGA 1056
QY 749 GGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTCAGCAGGAGTTTAAACAGAGCCT 808
DB 1057 GAGTAAGAGAAATAATATCAATGCTGTTCCGATAGAACAGGAGTCTCTCACTGAGGCT 1116
QY 809 TGCCAGTTGGCCTCATTTGGAATGCAATTTGATGAAACAGTACATTTGAGTTTGTAG 868
DB 1117 TGCTGTGTAAGCTCATTTGGGATGAAATTTGCACTCTAATGAAGCAATACATTTGAGTTTGTG 1176
QY 869 CTGACAGATTACTTTGGAACTTTGGATTCTCAAGGTTTTCAGGCAAGAAAACTCTTTG 928
DB 1177 CAGACAGATTATGCTGGAACTGGGTTTGTAGCAAGGTTTTCAGAGTAGAGAACCCATTG 1236
QY 929 ATTTTATGGAACAATTTCTTTAGAGGAAAAACAAATTTCTTTTGAAAAACGAGTTTCAG 988
DB 1237 ACTTTATGAGAATAATTTTCACTGGAAGGAAAGACTAACTTTCTTTTGAAGAGAGTAGCG 1296
QY 989 AGTATCAGCGTTTGGCATTTATGGCAGAAACACAGATAACGCTCTTCCACCTTGGATCGAC 1048

Db 1297 ACTATCAGAGGATGGAGTGTGATGTCAGTCCACAGAGATCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 12
ABL65859
ID ABL65859 standard; DNA; 2500 BP.
XX
AC ABL65859;
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:4196.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 29-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.

XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 4196; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;
Query Match 55.1%; Score 580; DB 6; Length 2500;
Best Local Similarity 75.1%; Pred. No. 1.4e-154;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
QY 89 ATGAAGAGCCACTCTCTAAGAAGAGTCTCGCGGTTTGTTCATCTTCCAAATCCAGTACC 148
Db 397 AGGATGAGCGCTGCTGAGAGAAAACCCCGCGCTTGTTCATCTTCCCATCGAGTACC 456
QY 149 CTGATATTTGGAATGTATTAACAGGCACAGCTTCTTCGCGGTTTGTTCATCTTCCAAATCCAGTACC 208
Db 457 ATGATATCTGGCAGATGTATAGAAGGAGGAGGCTTCTTTGGACCGCAGGAGGTTG 516
QY 209 ACTTATCAAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTTCATCT 268
Db 517 ACCTCTCAAGGACATTGACACTGGGAATCCCTGAACCCGAGGAGATATTTTATAT 576
QY 269 CTCACATCTTAGCCTTTTGTGAGCCAGTGAATGGAATGTAATGAAATTTGTGGAGC 328
Db 577 CCCATGTTCTGGCTTTCTTTGACGAAGCGATGGCATAGTAATGAAACTTGTGGAGC 636
QY 329 GCTTTAGTCAGGAGTGCAGGTTCCAGAGCTCGCTGTTCTATGGCTTCCAAATCTCA 388
Db 637 GATTTAGCCAAGAAGTTTCAGATTCAGAGCCCGCTGTTCTATGGCTTCCAAATGCCA 696
QY 389 TCGAGAAATGTTCACTCAGAGATGTACAGTTTGTGTGATAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAAACATACATTCTGAAATGTATAGTCTTTTATTTGACACTTACATAAAGATCCCA 756
QY 449 AGAAAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCCCTATGTTTAAAGAAAAAGCAG 508
Db 757 AAGAAAAGGGAATTTCTCTTCAATGCCATTGAAAGGATGCCCTTGTGTCAAGAAGAAGCAG 816
QY 509 ATTGGCCCTTGGCATGGATAGCAGATAGAAAATCTACTTTTGGGAAAGAGTGTGGCCT 568
Db 817 ACTGGGCCCTTGGCCTGGATTTGGGGACAAGAGAGGCTACCTATGTTGTAACGTTGTAGCCT 876
QY 569 TTGCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTGTGCTGTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGAGTGGAAAGGCAATTTCTTTTCCGGTCTTTTGTGCTGATATTTCTGGCTCAAGA 936

QY 629 AGAGAGGCTTATGCCAGGACTCAGCTTTTCCAAATGAACTCATCAGCAGAGATGAGGAC 688
 Db 937 AACGAGGAGTATGCTGCGCTTCAACATTTCTAATGAATCTTATAGCAGAGATGAGGTT 996
 QY 689 TTCCTGTGACTTGTCTGCTGATGTTCCAAATACCTTAGTAATAGCCTTCAGAGAAA 748
 Db 997 TACACTGTGATTTGCTTGCCTGATGTTCCAAACCTGCTGTACACAAACCATCGAGGAGA 1056
 QY 749 GGGTCAGGAGATCATTTGATGCTGTCAAAATTCAGCAGAGATTTTAAACAGAGCCT 808
 Db 1057 GAGTAAGAGAAAATAATTCAATGCTGTTCCGATAGAACAGAGATTCCTCACTGAGCCT 1116
 QY 809 TCCAGTGTGCTCCTATGGAATGAAATGCAATTTGATGAAACAGTACATGATTTGTAG 868
 Db 1117 TCCCTGTGAGCTCATTTGGATGGAATGCACTTAAATGAAGCAATACATGATTTGTGG 1176
 QY 869 CTGACAGATTAATCTGGAACCTTGATTTCTCAAGGTTTTCAGGCAGAAAATCCTTTTG 928
 Db 1177 CAGACAGACTTATGCTGNACTGGTGTGTTAGCAAGGTTTTCAGATGAGAACCATTTG 1236
 QY 929 ATTTATGGAACATTTCTTTAGAGGAAAACAAATTTCTTTGAGAAACGAGTTTTCAG 988
 Db 1237 ACTTTATGAGAAATATTTCACTGGAAGGAAAGACTAACTTTCTTTGAGAGAGAGTAGGCG 1296
 QY 989 AGTATCAGGCTTTGAGTATGCGAGAAACACAGATTAACGTTTCACTTTGGATGAG 1048
 Db 1297 AGTATCAGAGGATGGAGTGTCAAGTCCAACAGAGAAATCTTTTACCTTTGGATGCTG 1356
 QY 1049 ATTT 1052
 Db 1357 ACTT 1360

RESULT 13

ID ABX10335 standard; DNA; 2500 BP.
 XX
 AC ABX10335;
 XX
 DT 28-JAN-2003 (first entry)
 XX
 DE DNA encoding protein differentially regulated in prostate cancer #4.
 XX
 KW Prostate cancer; gene expression; differential regulation;
 KW molecular marker; drug target; cancer detection; cancer diagnosis;
 KW cancer staging; cancer grading; cancer assessing; cancer monitoring;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200281638-A2.
 XX
 PD 17-OCT-2002.
 XX
 PF 08-APR-2002; 2002WO-US010824.
 XX
 PR 06-APR-2001; 2001US-0281731P.
 PR 06-APR-2001; 2001US-0281732P.
 XX
 PA (ORIG-) ORIGENE TECHNOLOGIES INC.
 XX
 PI Sun Z, Jay G;
 XX
 DR WPI; 2003-058520/05.
 DR P-PSDB; ABU07433.
 XX
 PT Novel genes which are differentially regulated in prostate cancer, useful
 PT for diagnosing prostate cancer in prostate tissue sample and assessing
 PT therapeutic or preventive intervention in prostate cancer patients.
 XX
 PS Claim 1; Page 71-72; 416pp; English.

CC The invention describes genes (I) which are differentially regulated in
 CC prostate cancer. (I) Is useful for diagnosing a prostate cancer in a
 CC sample comprising prostate tissue, which involves determining the number
 CC of target genes which are differentially-regulated in the sample, where
 CC the number is indicative of the probability that the sample comprises
 CC prostate cancer. (I) Is useful for assessing a therapeutic or preventive
 CC intervention in a subject having a prostate cancer, which involves
 CC determining the expression levels in a sample comprising prostate tissue
 CC of target genes which are differentially-regulated in prostate cancer.
 CC Preferably, the expression levels of at least 10 genes are determined.
 CC (I) is also useful for identifying agents that modulate a biological
 CC activity of a polypeptide differentially-regulated in prostate cancer
 CC cells, which involves contacting a polypeptide differentially-regulated
 CC in prostate cancer cells with a test agent under conditions effective for
 CC the test agent to modulate a biological activity of the polypeptide, and
 CC determining whether the test agent modulates the biological activity. (I)
 CC is useful as molecular markers, as drug targets, and for detecting,
 CC diagnosing, staging, grading, assessing, monitoring, prognosticating,
 CC preventing or treating, determining predisposition to diseases and
 CC products especially relating to prostate cancer. (I) and its expression
 CC products are used in the diagnostic test to assay for presence of cancer
 CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
 CC blood etc. (I) is useful for assessing cancer e.g., to determine the type
 CC of cancer, its stage of development, the nature of genetic defect, etc.
 CC The polypeptide encoded by (I) can be used as target for therapy or drug
 CC discovery. (I) can also be used for expressing the polypeptide and thus
 CC for searching specific binding partners of the polypeptide. (I) is useful
 CC in therapeutic applications to treat prostate cancer. The identification
 CC of specific genes, and groups of genes, expressed in pathways
 CC physiologically relevant to prostate cancer permits the definition of
 CC functional and disease pathways and the delineation of targets in these
 CC pathways which are useful in diagnostic, therapeutic, and clinical
 CC applications. This sequence encodes a protein differentially regulated in
 CC prostate cancer
 XX
 SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;

Query Match 55.1%; Score 580; DB 8; Length 2500;
 Best Local Similarity 75.1%; Pred. No. 1.4e-154;
 Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
 QY 89 ATGAGAGGCACTCTTAAGAAAGATTCCTCGCGGTTTGTTCATCTTTCCATCCAGTACC 148
 Db 397 AGATGAGCGCTGCTGAGAGAAAACCCCGCGCTTTGTTCATCTTTCCCATCCAGTACC 456
 QY 149 CTGATATTTGGAAAATGTATAAACAGCAGCAGGCTTCTCTTGACAGCAGAGAGGTTG 208
 Db 457 ATGATATCTGGCAGATGTATAGAACGACAGAGCTTCCCTTTGGACCGCGAGAGGTTG 516
 QY 209 ACTTATCAAAGGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTACTTCATCT 268
 Db 517 ACCTCTCCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
 QY 269 CTCACATCTTAGCCTTTTTCAGCGCAGTGATGGAATTTAAATGAAAATTTGTTGAGC 328
 Db 577 CCCATGTTCTGCGCTTTCTTTGCAAGCAAGCATGGCATAGTAAATGAAAATCTTGTGGAGC 636
 QY 329 GCTTTAGTCAGGAGGTGCAAGTTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTTCTCA 388
 Db 637 GATTTAGCCCAAGAGTTTCAGATTACAGAGCCGCTGTTTCTATGGCTTTCAAATTTGCCA 696
 QY 389 TCGAGAATTTCACTCAGAGATGTACAGTTTGTGTATAGACACTTACATCAGAGATCCCA 448
 Db 697 TGGAAAACATACATCTGAAAATGTATAGTCTTCTTATTGACACTTACATAAAGATCCCA 756
 QY 449 AGAAAAGGGAATTTTATTATTAATGCAATTTGAAACCATGCCCCTATGTTAAGAAAAAGCAG 508
 Db 757 AAGAAAAGGGAATTTCTCTCAATGCCATTGAAAACGATGCCCTTGTGTCAAGAAAGGAGCAG 816
 QY 509 ATTGCGCCTTGGATGGATAGCAGATAGAAAATCTACTTTTGGGAAAGAGTGTGTGCGCT 568
 Db 817 ACTGGCCTTGGCTGGATTTGGGACAAAGAGGCTTACCTATGTGTGAACGTGTGTAGCCT 876

QY 629 AGAGAGGCTTATGCCAGGACTCACATTTTCCAAATGAATCATCAGCAGAGATGAAGGAC 688
 Db
 QY 937 AACGAGGACTGATGCGCTGCGCTCACATTTTCTAATGAATCTATTAGCAGAGATGAGGTT 996
 Db
 QY 689 TTCACTGTGACTTGTGCTGCTGATGTTCAATCTATTAGTAAATTAAGCCTTCAGAGAAA 748
 Db
 QY 997 TACACTGTGATTTTGTGCTGCTGATGTTCAACACCTGGTACACAAACCATCGGAGGAGA 1056
 Db
 QY 749 GGTCTAGGAGATCATTTGATGCTGCTCAAAATGAGCAGGAGTTTAAACAGAGCCT 808
 Db
 QY 1057 GAGTAAGAGAAATAATTTATCAATGCTGTTCCGATAGAACAGGAGTTCCCTCAGGAGCCT 1116
 Db
 QY 809 TGGCAGTGGCCCTCAITTTGAATGAATGCAATTTTGAATGAACAGTACATTTGATGTTGTAG 868
 Db
 QY 1117 TGCTGTGAAGCTCATTTGGATGAATTTGCACTCTAATGAAGCAATACATTTGATTTGTGG 1176
 Db
 QY 869 CTGACAGATTTACTTTGGAACTTTGATTTCTCAAGGTTTTTCAGGCGAGAAAATCTCTTTG 928
 Db
 QY 1177 CAGACAGACTTATGCTGGAACCTGGTTTGTAGCAAGGTTTTTCAGAGTAGAGAACCCATTG 1236
 Db
 QY 929 ATTTTATGGAACATTTCTTTTGAAGGAGAAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
 Db
 QY 1237 ACTTTATGGAGATATTTTCACTGGAAGAAAGACTAACTTTCTTTTGAGAGAGATAGGCG 1296
 Db
 QY 989 AGTATCAGGCTTTTGCAGTTTATGCGAGAAACACAGATACGTCCTTACCTTCGATGCGAG 1048
 Db
 QY 1297 AGTATCAGAGATGGAGATGATCAAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTG 1356
 Db

RESULT 15
 ADN04443
 ID ADN04443 standard; cDNA; 2500 BP.
 AC ADN04443;

01-JUL-2004 (first entry)

Antipsoriatic cDNA sequence #423.

de; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.

Homo sapiens.

WO2004028479-A2.

08-APR-2004.

25-SEP-2003; 2003WO-US030907.

25-SEP-2002; 2002US-0414006P.

(GETH) GENENTECH INC.

Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PW, Wood WI;
 Wu TD;

WPI; 2004-305105/28.

P-PSDB; ADN04444.

New PRO nucleic acid or polypeptide, useful for preparing a
 pharmaceutical composition for diagnosing or treating psoriasis in a
 mammal.

Claim 1; SEQ ID NO 837; 3069pp; English.

The invention relates to novel polynucleotide and polypeptides for
 treating psoriasis or a sequence having at least 80% identity to the
 above sequences. The nucleic acid is useful for preparing a composition
 for diagnosing or treating psoriasis in a mammal. This sequence

CC corresponds to one of the polynucleotides of the invention.
 XX
 SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;

Query Match 55.1%; Score 580; DB 12; Length 2500;
 Best Local Similarity 75.1%; Pred. No. 1.4e-154;
 Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCACATCTTAAGAAAGAGTTCTCGCCGGTTTGTCTTCCAAATCCAGTACC 148
 Db
 QY 397 AGGATGAGCGCTGCTGAGAGAAAACCCCGCCGCTTTGTCTTCCCATCGAGTACC 456
 Db
 QY 149 CTGATATTTGCAAAAATGATAAAGCAGCAGAGCTTCTCTGTCGACGACAGAGGTTG 208
 Db
 QY 457 ATGATATCTGCGAGATGATTAAGAGCAGAGGCTTCTTTTGGACCCCGAGGAGTTG 516
 Db
 QY 209 ACTTATCAAAAGGATCTCCCTCCTCAGTGAACCAAGCTTAAAGCAGATGAGAAGTACTTTCATCT 268
 Db
 QY 517 ACCTCTCCAAGGACATTCAGCACTGGGAATCCCTGMAACCCGAGAGAGATATTTTATAT 576
 Db
 QY 269 CTCACATCTTAGCCCTTTTTCAGCCAGTGAAGAAATGTAATGAAAAATTTGTGGAGC 328
 Db
 QY 577 CCCATGTTCTGGCTTTCTTTTCAGCAAGCGATGTCATAGTAAATGAAAACTTGGTGGAGC 636
 Db
 QY 329 GCTTTAGTCAGAGGTTCAGGTTCCAGAGCTCGCTGTTCTATGCTTTCAAAATCTCA 388
 Db
 QY 637 GATTTAGCCAAAGATTCAGATTACAGAGCCCGCTGTTCTATGCTTCCAAATTCGCA 696
 Db
 QY 389 TCGAGAATGTTCACTCAGAGATGTACAGTTTGTGTAGTAGACACTTACATCAGAGATCCCA 448
 Db
 QY 697 TGAAGAAACATACATTTCTGAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 756
 Db
 QY 449 AGAAAAGGGAATTTTATTTAATGCAATGAAACCATGCCCTATGTTAAGAAAAAAGCAG 508
 Db
 QY 757 AAGAAAAGGGAATTTCTCTTCAATGCCATTGAAACGATGCTTGTGTCAAGAAGAGGCGAG 816
 Db
 QY 509 ATTGGGCTTGCATGATGATAGAGATAAGAAATCTACTTTTGGGGAAGAGTGGTGGCT 568
 Db
 QY 817 ACTGGGCTTGGCTGGATTTGGGGAACAAAGAGGCTACTATGTTGTAACGCTGTTGAGCCT 876
 Db
 QY 569 TTGCTGTGTGAGAGGAGTTTCTCTCAGGATCTTTTGTGCTGTATATCTTGGCTAAAGA 628
 Db
 QY 877 TTGCTGAGTGGAGGCAATTTCTTTCCGGTCTTTTGGCTGCAATTTCTGGCTCAAGA 936
 Db
 QY 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGGAC 688
 Db
 QY 937 AACGAGGACTGATGCTGCGCTCACATTTTCTAATGAACCTTATAGCAGAGATGAGGTT 996
 Db
 QY 689 TTCACTGTGACTTTGCTTGGCTGATGTTCAATCTATTAGTAAATAAGCCTTCAGAGAAA 748
 Db
 QY 997 TACACTGTGATTTTGTGCTTGGCTTCAACACACCTGGTACACAAACCATCGGAGGAGA 1056
 Db
 QY 749 GGTCTAGGAGATCATTTGTTGATGCTGCTCAAAATGAGCAGGAGTTTAAACAGAGCCT 808
 Db
 QY 1057 GAGTAAGAGAAATAATTTATCAATGCTGTTCCGATAGAACAGGAGTTCTCTCAGTGGCCT 1116
 Db
 QY 809 TGGCAGTGGCCCTCATTTGGAATGAATTTGCAATTTTGTATGAAACAGTACATTTGATGTTGTAG 868
 Db
 QY 1117 TGCTGTGAAGCTCATTTGGATGAATTTGCACTCTAATGAAGCAATACATTTGATTTGTGG 1176
 Db
 QY 869 CTGACAGATTTACTTTGGAACTTTGATTTCTCAAGGTTTTTCAGGCGAGAAAATCTCTTTG 928
 Db
 QY 1177 CAGACAGACTTATGCTGGAACCTGGTTTGTAGCAAGGTTTTTCAGAGTAGAGAACCCATTG 1236
 Db
 QY 929 ATTTTATGGAACATTTCTTTTGAAGGAGAAAACAAATTTCTTTTGAAGAACGAGTTTCAG 988
 Db
 QY 1237 ACTTTATGGAGATATTTTCACTGGAAGAAAGACTTAACTTTCTTTTGAAGAGAGATAGGCG 1296
 Db
 QY 989 AGTATCAGGCTTTTTCAGTTTATGCGAGAAACACAGATAACGCTTCCACCTTGGATGCGAG 1048
 Db
 QY 1297 AGTATCAGAGATGGAGATGATGCAAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTG 1356
 Db

1049 ATTT 1052

Db 1357 ACTT 1360

Search completed: October 30, 2005, 00:02:46
Job time : 442.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2005, 23:30:35 ; Search time 146 seconds
(without alignments)
11801.366 Million cell updates/sec

Title: US-10-698-228-12

Perfect score: 1053

Sequence: 1 atggggcgaccggaaggcc.....tcacctggatgcagatttt 1053

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/prodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/prodata/1/ina/PCTUS COMB.seq.*
- 6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1053	100.0	1053	4	US-10-019-733-12
2	1053	100.0	1081	4	US-10-019-733-4
3	1051.4	99.8	1053	4	US-10-019-733-2
4	1051.4	99.8	4955	4	US-10-019-733-3
5	580.6	55.1	2500	4	US-09-962-665-9
6	580.6	55.1	2500	4	US-09-963-333-9
7	580.6	55.1	2500	4	US-09-962-677-9
8	580	55.1	2500	4	US-09-023-655-1370
9	580	55.1	2500	4	US-09-949-016-145
10	578.4	54.9	2479	3	US-09-949-016-2025
11	439.4	41.7	481	3	US-08-905-223-125
12	429.8	40.8	14176	1	US-08-307-499-1
13	429.8	40.8	14176	1	US-08-307-499-14
14	429.8	40.8	14176	3	US-09-299-268-1
15	429.8	40.8	14176	3	US-09-299-268-14
16	384.8	36.5	1236	4	US-09-248-796A-3914
17	382.8	36.4	825	3	US-08-307-499-40
18	382.8	36.4	825	3	US-09-299-268-40
19	339.4	32.2	1112	4	US-08-136-743B-1
20	317.6	30.2	1083	4	US-09-248-796A-3915
21	216.6	30.1	419	3	US-08-905-223-129
22	260.8	24.8	656	3	US-09-328-111-791
23	195.2	18.5	449	4	US-09-270-767-14670
24	169.2	16.1	366	3	US-08-307-499-13
25	169.2	16.1	366	3	US-09-299-268-13
26	141.4	13.4	373	4	US-09-401-064-337
27	139.8	13.3	373	4	US-09-401-064-319

c	28	115.2	10.9	366	4	US-09-401-064-307	Sequence 307, App
	29	112.4	10.7	297	4	US-09-313-294A-4435	Sequence 4435, Ap
	30	93.8	8.9	11820	4	US-09-949-016-11887	Sequence 11887, A
	31	93.8	8.9	11826	4	US-09-949-016-13767	Sequence 13767, A
	32	92	8.7	601	4	US-09-949-016-20615	Sequence 20615, A
	33	92	8.7	601	4	US-09-949-016-69028	Sequence 69028, A
	34	76	7.2	276	4	US-09-313-294A-3437	Sequence 3437, Ap
	35	71.8	6.8	304	4	US-09-313-294A-7143	Sequence 7143, Ap
	36	68	6.5	276	4	US-09-313-294A-4638	Sequence 4638, Ap
	37	61	5.8	1053	4	US-09-902-540-2666	Sequence 2666, Ap
c	38	61	5.8	13706	4	US-09-902-540-1124	Sequence 1124, Ap
	39	60.8	5.8	1141	4	US-09-806-708B-22	Sequence 22, Appl
	40	60	5.7	601	4	US-09-949-016-69031	Sequence 69031, A
	41	54.4	5.2	288	4	US-09-313-294A-6108	Sequence 6108, Ap
	42	53.8	5.1	276	4	US-09-313-294A-3244	Sequence 3244, Ap
	43	50	4.7	1230025	4	US-09-198-452A-1	Sequence 1, Appli
	44	50	4.7	1230230	4	US-09-438-185A-1	Sequence 1, Appli
c	45	47	4.5	124884	4	US-09-661-596A-76	Sequence 76, Appl

ALIGNMENTS

RESULT 1

US-10-019-733-12

; Sequence 12, Application US/10019733

; Patent No. 6682917

; GENERAL INFORMATION:

; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA

; TITLE OF INVENTION: New Protein and its DNA

; FILE REFERENCE: 2619W00P

; CURRENT APPLICATION NUMBER: US/10/019,733

; PRIOR FILING DATE: 2001-12-28

; PRIOR APPLICATION NUMBER: JP 11-181131

; PRIOR FILING DATE: 1999-06-28

; PRIOR APPLICATION NUMBER: JP 11-192391

; PRIOR FILING DATE: 1999-07-06

; PRIOR APPLICATION NUMBER: JP 2000-017770

; PRIOR FILING DATE: 2000-01-21

; NUMBER OF SEQ ID NOS: 14

; SEQ ID NO 12

; LENGTH: 1053

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION:

US-10-019-733-12

Query Match 100.0%; Score 1053; DB 4; Length 1053;
Best Local Similarity 100.0%; Pred. No. 6.8e-309;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGGGGCACC	CGGAAAGCGCGGCTGGATCAGGATGAGATCATCTTCA	60
Db	1	ATGGGGCACC	CGGAAAGCGCGGCTGGATCAGGATGAGATCATCTTCA	60
Qy	61	GACACCAAC	GAAAGTGAATATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC	120
Db	61	GACACCAAC	GAAAGTGAATATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC	120
Qy	121	CGGTTTGT	CATCTTTTCAATCCAGTACCTTGATTTTGGAAAATGTATATAACAGGCACAG	180
Db	121	CGGTTTGT	CATCTTTTCAATCCAGTACCTTGATTTTGGAAAATGTATATAACAGGCACAG	180
Qy	181	GCTTCTCT	TGGACACAGAGAGTTGACTTATCAAGAGATCTCCCTCACTGGAACAG	240
Db	181	GCTTCTCT	TGGACACAGAGAGTTGACTTATCAAGAGATCTCCCTCACTGGAACAG	240
Qy	241	CTTAAAGC	AGATGAGAAGTACTTCTCATCTTACACATCTTTAGCCCTTTTTCGAGCCAGTGAT	300
Db	241	CTTAAAGC	AGATGAGAAGTACTTCTCATCTTACACATCTTTAGCCCTTTTTCGAGCCAGTGAT	300
Qy	301	GGAATTGT	ATAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGAGGTTCCAGAGGCT	360

Db 301 GGAATGTAATAAGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
Qy 361 CGCTGTTCTATGGCTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTCTATGGCTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATCAATTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATCAATTGAA 480
Qy 481 ACCATGCCCCATGTTAAGAAAAAGCAGATTGGGCTTGCATGATAGCAGATGAAAA 540
Db 481 ACCATGCCCCATGTTAAGAAAAAGCAGATTGGGCTTGCATGATAGCAGATGAAAA 540
Qy 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAAGGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAAGGAGTTTCTTCTCAGGA 600
Qy 601 TCTTTTGTGCTATATTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC 660
Db 601 TCTTTTGTGCTATATTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC 660
Qy 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 720
Db 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 720
Qy 721 TACTTAGTAATAAGCCCTCAGAAAGAGGTTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db 721 TACTTAGTAATAAGCCCTCAGAAAGAGGTTCAGGAGATCATTTGTTGATGCTGTCAA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGATGAGGTTCAGGAGATCATTTGTTGATG 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGATGAGGTTCAGGAGATCATTTGTTGATG 840
Qy 841 TTGATGAAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGGAATTCGATTTCTCA 900
Db 841 TTGATGAAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGGAATTCGATTTCTCA 900
Qy 901 AAGGTTTTTCAGGCGAAGAAATCCTTTGATTTTATGAAACATTTCTTTAGAGGAAAA 960
Db 901 AAGGTTTTTCAGGCGAAGAAATCCTTTGATTTTATGAAACATTTCTTTAGAGGAAAA 960
Qy 961 ACAAAATTTCTTTGAGAAAGAGTTTCAGAGTATCAGGTTTTGTCAGTTATGCGAGAAC 1020
Db 961 ACAAAATTTCTTTGAGAAAGAGTTTCAGAGTATCAGGTTTTGTCAGTTATGCGAGAAC 1020
Qy 1021 ACAGATAACGCTTCCACCTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGCTTCCACCTTGGATGAGATTTT 1053

RESULT 2

US-10-019-733-4
; Sequence 4, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens

; FEATURE:
; OTHER INFORMATION:
US-10-019-733-4
Query Match 100.0%; Score 1053; DB 4; Length 1081;
Best Local Similarity 100.0%; Pred. No. 6.9e-309;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGGGACCCCGAAAGCGCGGCGGGCTGGATCAGATCAGAGATCATCTTCA 60
Db 20 ATGGGGACCCCGAAAGCGCGGCGGGCTGGATCAGATCAGAGATCATCTTCA 79
Qy 61 GACACCAACGAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTCTCGC 120
Db 80 GACACCAACGAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTCTCGC 139
Qy 121 CGGTTTGTGCTATTTTCCAAATCCAGTACCTGATATTTGAAAAATGTATAAACAGGCACAG 180
Db 140 CGGTTTGTGCTATTTTCCAAATCCAGTACCTGATATTTGAAAAATGTATAAACAGGCACAG 199
Qy 181 GCTTCTTCTGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db 200 GCTTCTTCTGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 259
Qy 241 CTTAAAGCAGATGAGAAGTACTTCTCTCATCTTACAGCTTTTGTAGCCCTTTTTCAGCAGCTGAT 300
Db 260 CTTAAAGCAGATGAGAAGTACTTCTCTCATCTCTCATCTTACAGCTTTTTCAGCAGCTGAT 319
Qy 301 GGAATGTAATAAGAAATTTGGTGAGCGCTTTAGTCCAGAGGTGCAGGTTCCAGAGGCT 360
Db 320 GGAATGTAATAAGAAATTTGGTGAGCGCTTTAGTCCAGAGGTGCAGGTTCCAGAGGCT 379
Qy 361 CGCTGTTCTATGGCTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 420
Db 380 CGCTGTTCTATGGCTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 439
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 499
Qy 481 ACCATGCCCCATGTTAAGAAAAAGCAGATTGGGCTTGCAGATGAGATGAGAAAA 540
Db 500 ACCATGCCCCATGTTAAGAAAAAGCAGATTGGGCTTGCAGATGAGATGAGAAAA 559
Qy 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAAGAGGATTTCTTCTCAGGA 600
Db 560 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAAGAGGATTTCTTCTCAGGA 619
Qy 601 TCTTTTGTGCTATATTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC 660
Db 620 TCTTTTGTGCTATATTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC 679
Qy 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGATGTTCCAA 720
Db 680 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGATGTTCCAA 739
Qy 721 TACTTAGTAATAAGCCCTTCAAGAAAGGCTTCAAGAGATCATTTGATGCTGTCAAA 780
Db 740 TACTTAGTAATAAGCCCTTCAAGAAAGGCTTCAAGAGATCATTTGATGCTGTCAAA 799
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGATGAGGTTCAGTTGGAATGATTCGATT 840
Db 800 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGATGAGGTTCAGTTGGAATGATTCGATT 859
Qy 841 TTGATGAAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGGAATGATTCGATTCTCA 900
Db 860 TTGATGAAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGGAATGATTCGATTCTCA 919
Qy 901 AAGGTTTTTCAGGCGAAGAAATCCTTTGATTTTATGAAAAATTTCTTTAGAGGAAAA 960
Db 920 AAGGTTTTTCAGGCGAAGAAATCCTTTGATTTTATGAAAAATTTCTTTAGAGGAAAA 979
Qy 961 ACAAAATTTCTTTGAGAAAGAGTTTCAGAGTATCAGGTTTTGTCAGTTATGCGAGAAC 1020

Db 980 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGGCAGAAACC 1039
QY 1021 ACAGATAACGCTCTTACCTTGGATGACAGATTTT 1053
Db 1040 ACAGATAACGCTCTTACCTTGGATGACAGATTTT 1072

RESULT 3

US-10-019-733-2
; Sequence 2, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-2

Query Match 99.8%; Score 1051.4; DB 4; Length 1053;
Best Local Similarity 99.9%; Pred. No. 2.1e-308;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
Db 1 ATGGGCGACCCGGAAGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
QY 61 GACACCAACGAAAGTGAATAAGTCAATGAAGGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATAAGTCAATGAAGGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 121 CGGTTTGTTCATCTTCCAAATCCAGTACCTCGATATTTGGAAAAATGATATAACAGGCACAG 180
Db 121 CGGTTTGTTCATCTTCCAAATCCAGTACCTCGATATTTGGAAAAATGATATAACAGGCACAG 180
QY 181 GCTTCCTTCTGGACAGCAGAGAGGTTGACTTATCAAGAGTCTCCCTCACTGGAAACAAG 240
Db 181 GCTTCCTTCTGGACAGCAGAGAGGTTGACTTATCAAGAGTCTCCCTCACTGGAAACAAG 240
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCTCATCTTAGCCTTTTTCAGGCCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAAGTACTTCTCTCATCTTAGCCTTTTTCAGGCCAGTGAT 300
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
Db 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAAGTGTTCACATGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAAGTGTTCACATGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCTTCGATGGATAGCAGATAGAAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCTTCGATGGATAGCAGATAGAAAA 540

QY 541 TCTACTTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGATTTTCTTCTCAGGA 600
Db 541 TCTACTTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGATTTTCTTCTCAGGA 600
QY 601 TCTTTTGTGCTATATTCTGGCTAAAGAGAGAGGTCTTATGCGCAGACTCATCTTTTCC 660
Db 601 TCTTTTGTGCTATATTCTGGCTAAAGAGAGAGGTCTTATGCGCAGACTCATCTTTTCC 660
QY 661 AATGAACTCATCAGCAGAGATGAAGGACTTCACCTGTGACCTTGTGCTGCTGATGTTCCAA 720
Db 661 AATGAACTCATCAGCAGAGATGAAGGACTTCACCTGTGACCTTGTGCTGCTGATGTTCCAA 720
QY 721 TACTTAGTAATAAGCCCTTCAGAGAAAGGCTCAGGAGAGTCAATTGTTGATGCTGTCAAA 780
Db 721 TACTTAGTAATAAGCCCTTCAGAGAAAGGCTCAGGAGAGTCAATTGTTGATGCTGTCAAA 780
QY 781 ATTGAGCAGAGTTTTAAACAGAAAGCTTGCAGTGGCCCTCATTTGGAATGAATTCATT 840
Db 781 ATTGAGCAGAGTTTTAAACAGAAAGCTTGCAGTGGCCCTCATTTGGAATGAATTCATT 840
QY 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Db 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
QY 901 AAGGTTTTTTCAGGCAGAGAAATCCTTTTGTATTTATGAAAAACATTTCTTTAGAAAGAAA 960
Db 901 AAGGTTTTTTCAGGCAGAGAAATCCTTTTGTATTTATGAAAAACATTTCTTTAGAAAGAAA 960
QY 961 ACAAAATTTCTTTCAGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
Db 961 ACAAAATTTCTTTCAGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
QY 1021 ACAGATAACGCTCTTACCTTGGATGACAGATTTT 1053
Db 1021 ACAGATAACGCTCTTACCTTGGATGACAGATTTT 1053

RESULT 4

US-10-019-733-3
; Sequence 3, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 3
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-3

Query Match 99.8%; Score 1051.4; DB 4; Length 4955;
Best Local Similarity 99.9%; Pred. No. 5.1e-308;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
Db 245 ATGGGCGACCCGGAAGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 305 GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364

```
QY 121 CGGTTTGTTCATCTTTCCAAATCAGTACCTGTGATATTGGAAAAATGTATAAAGGACAG 180
Db 365 CGGTTTGTCAUCTTTCCAAATCAGTACCTGTGATATTGGAAAAATGTATAAAGGACAG 424
QY 181 GCTTCCTTCTGGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db 425 GCTTCCTTCTGGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 484
QY 241 CTTAAAGCAGATGAGAGTACTTCTCTCTCAATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db 485 CTTAAAGCAGATGAGAGTACTTCTCTCTCAATCTTAGCCCTTTTTCAGCCAGTGAT 544
QY 301 GGAATTGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCCAGAGGCT 360
Db 545 GGAATTGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCCAGAGGCT 604
QY 361 CGCTGTCTTATGGCTTTCAAAATCTCATCGAAGTGTTCACCTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTCTTATGGCTTTCAAAATCTCATCGAAGTGTTCACCTCAGAGATGTACAGTTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTAATGCAATTGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTAATGCAATTGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTCGGCCTTCGGATGGATAGCAGATGAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATTCGGCCTTCGGATGGATAGCAGATGAAA 784
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTTGTCTGTAGAGGAGTTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGTGGCTTTTGTCTGTAGAGGAGTTTTCTTCTCAGGA 844
QY 601 TCTTTTGTGCTATATCTTGGCTTAAGAAAGAGGTCCTTATGCCAGGACTCATTTTTC 660
Db 845 TCTTTTGTGCTATATCTTGGCTTAAGAAAGAGGTCCTTATGCCAGGACTCATTTTTC 904
QY 661 AATGAATCATCAGCAGAGATGAAGACCTTCACTGTGACTTTGCTTGCCTGTGATGTTCCAA 720
Db 905 AATGAATCATCAGCAGAGATGAAGACCTTCACTGTGACTTTGCTTGCCTGTGATGTTCCAA 964
QY 721 TACTTAGTAAATGAAGCCTTTCAGAGAAAGGGTCAGGAGATCAATTTGTGATGCTGCAAA 780
Db 965 TACTTAGTAAATGAAGCCTTTCAGAGAAAGGGTCAGGAGATCAATTTGTGATGCTGCAAA 1024
QY 781 ATTGACGAGGATTTTAAAGAAAGCCTTGCAGATTTGGCCTCATTTGGAATGAATTGCAIT 840
Db 1025 ATTGACGAGGATTTTAAAGAAAGCCTTGCAGATTTGGCCTCATTTGGAATGAATTGCAIT 1084
QY 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTTACTTGTGGAATTTGATTTCTCA 900
Db 1085 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTTACTTGTGGAATTTGATTTCTCA 1144
QY 901 AAGGTTTTTTCAGCAGAAATCCTTTTGATTTTATGAAAAACATTTCTTTAGAAAGAAAA 960
Db 1145 AAGGTTTTTTCAGCAGAAATCCTTTTGATTTTATGAAAAACATTTCTTTAGAAAGAAAA 1204
QY 961 ACAATTTCTTGGAGAAAGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
Db 1205 ACAATTTCTTGGAGAAAGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1264
QY 1021 ACAGATAAGCTTTCACCTTGGATGCAGATTTT 1053
Db 1265 ACAGATAAGCTTTCACCTTGGATGCAGATTTT 1297
```

RESULT 5

US-09-962-665-9

; Sequence 9, Application US/09962665

; Patent No. 6537759

; GENERAL INFORMATION:

; APPLICANT: Stanton, Jr., Vincent P.

; TITLE OF INVENTION: POLYLPOLYGLUTAMATE SYNTHETASE GENE SEQUENCE

```
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING THE
; TITLE OF INVENTION: TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015004
; CURRENT APPLICATION NUMBER: US/09/962,665
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128, 1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc_feature
; LOCATION: 1636, 1738, 2259
; OTHER INFORMATION: n = c or t
; US-09-962-665-9
```

Query Match

55.1%; Score 580.6; DB 4; Length 2500;

Best Local Similarity 75.1%; Pred. No. 2,2e-165;

Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

```
QY 89 ATGAAGGACCTCTTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTCCATCCAGTACC 148
Db 397 AGGATGAGCGCTGCTGAGAGAAAAACCCCGCGCTTGTTCATCTTCCCATCGAGTACC 456
QY 149 CTGATATTTGAAAAATGTATAAAGCAGCAGAGCTTCTCTTGACACAGCAGAGGTTG 208
Db 457 ATGATATCTGGCAGATGTATAAGAGGCGAGAGCTTCTTTTGACCCCGAGGAGTTG 516
QY 209 ACTTATCAAGAGTATCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAGTACTTCACT 268
Db 517 ACCTCTCNAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
QY 269 CTCACATCTTAGCTTTTTCAGCCAGTATGGAATGTAAATGAAAAATTTGGTGAGC 328
Db 577 CCATGTTCTGGCTTTCTTTGCGAGCAAGCATGGCATAGTAAATGAAAAATTTGGTGAGC 636
QY 329 GCTTTTGTAGTCAGGAGGTGCGAGTTCCAGAGCTCGCTGTCTTCTATGGCTTTTCAAAATTTCTCA 388
Db 637 GATTAGCCNAGAGTTTCAGATTACAGAACCCGCTGTCTTCTATGGCTTCCAAATTTGCCA 696
QY 389 TCAGAGATGTTCACTCAGAGATGTACAGTTTGTGTATAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAAAACATACATTCTGAAATGTATAGTCTTCTTATTGACACTTACATAAAGATCCCA 756
QY 449 AGAAAAAGGAATTTTATTTAATGCAATTCGAATTCGAACCATGCCCTATGTTAAGAAAAAGCAG 508
Db 757 AAGAAAAAGGAATTTTCTCTTCAATGCCATTTGAAACGATGCCCTTGTGTCAAGAAAGAGGCG 816
QY 509 ATTGGGCTTGGATGGATGAGATAGAAAAATCTACTTTTGGGAAAGAGTGTGTGGCT 568
```

```
Db 817 ACTGGCCCTTGGCGCTGGATTGGGACAAAGAGCTACCTATGCTGAACGCTGTGTAGCCT 876
Qy 569 TTGCTGCTCTAGAAGAGTTTCTCTCAGGATCTTTTCTGCTGCTATATTTCTGCTAAAGA 628
Db 877 TTGCTGCTGAGGAGCAATTTCTTTCCGGTCTTTTGGCTGCTGATATTTCTGCTCAAGA 936
Qy 629 AGAGAGGTCTTATGCGAGGACTCAGTTTCCAACTCAATCAATCAATCAATCAATCAATCA 688
Db 937 AACGAGGACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 996
Qy 689 TTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 748
Db 997 TACACTGTGATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1056
Qy 749 GGGTCAGGAGATCATTTGTTGATGCTGCTCAAAATTCAGCAGGAGTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATAATTAATCAATGCTGTTCCGATAGAACAGGAGTTCCTCACTGAGCCT 1116
Qy 809 TGCCAGTTGGCCTCATTTGGAATGAATTCATTTTGAATGAACAGTACATTTGATTTGTAG 868
Db 1117 TGCTGTGAGCTCATTTGGATGAATTCGACTCTAATGAAGCAATACATTTGATTTGTGG 1176
Qy 869 CTGACAGATTACTTTGTGGAATTTGGATTCTCAAGGTTTTCAGGCAGAAATTCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACCTGGGTTTTCAGGAGGTTTTCAGAGTAGAACCCATTG 1236
Qy 929 ATTTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACAGATTTCAG 988
Db 1237 ACTTTATGGAATAATTTCACTGGAAGGAAAGACTAACTTTCTTTGAGAGAGATAGGCG 1296
Qy 989 AGTATCAGGCTTTTGCAGTTATGCGAGAAACACAGATAACGCTTCTCACTTTGGATGCG 1048
Db 1297 AGTATCAGAGATGAGGATGATGTCAGTCCAAACAGAGAAATTTCTTTTACCTTTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360
```

RESULT 6

```
US-09-963-333-9
; Sequence 9, Application US/09963333
; Patent No. 6664062
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE SYNTHASE GENE SEQUENCE VARIANCES
; TITLE OF INVENTION: HAVING UTILITY IN DETERMINING THE TREATMENT
; TITLE OF INVENTION: OF DISEASE
; FILE REFERENCE: 11926-015002
; CURRENT APPLICATION NUMBER: US/09/963,333
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128, 1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
```

```
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc_feature
; LOCATION: 1636, 1738, 2259
; OTHER INFORMATION: n = c or t
; US-09-963-333-9
```

```
Query Match 55.1%; Score 580.6; DB 4; Length 2500;
Best Local Similarity 75.1%; Pred. No. 2.2e-165;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
```

```
Qy 89 ATGAAGAGGCACCTCCCTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTTCCAAATCCAGTACC 148
Db 397 AGGATGAGCGCTGCTGAGAGAAAACCCCGCGCTTTGTCACTTTCCCATCGAGTACC 456
Qy 149 CTGATATTTGGAATAATGATAAAGCAGCAGCGTTCCTTCTGGACAGCAAGAGGTTG 208
Db 457 ATGATATCTGGCAGATGATAAAGAGCAGAGGCTTCTTTTGGACCGCGAGGAGTTG 516
Qy 209 ACTTATCAAGGATCTCCCTCACTGGAACAGCTTAAAGCAGATGAGAGTACTTCACT 268
Db 517 ACCTCTCAAGGACATTCAGACCTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
Qy 269 CTCACATCTTAGCCTTTTTCAGCCAGCCAGTGCATGGAATTTGAAATGAAAAATTTGGTGAGC 328
Db 577 CCCATGTTCTGGCTTTCTTTGAGCAAGCATGGCATAGTAATGAAACTTTGGTGAGC 636
Qy 329 GCTTTAGTCAGGAGGTGAGGTTCCAGAGCTCGCTGTTTCTATGCTTTCAAAATCTCA 388
Db 637 GATTTAGCCAAAGATTTCAGATTACAGAACCCGCTGTTTCTATGCTTCCAAATTCGA 696
Qy 389 TCGAGNAATGTTCACTCAGAGATGTACAGTTTGTGTGATAGACACTTACATCAGAGATCCA 448
Db 697 TGGAAAAACATACATTTCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCA 756
Qy 449 AGAAAGAGGAAATTTTATTTAATGCAATTTGAAACCATGCCCTATGTTTAAAGAAAAACAG 508
Db 757 AGAAAGAGGAAATTTCTTCAATGCCATGCAACGATGCCCTTGTGTCAAGAAAGAGCAG 816
Qy 509 ATTGGCCCTTGGCATAGCAGATAGCAAAATCTACTTTTGGGAAAGAGTGTGTCCT 568
Db 817 ACTGGCCCTTGGCTGGAATGGGACAAAGAGGCTACCTATGTTGAACGTTGTAGCCT 876
Qy 569 TTGCTGCTGTAGAGAGGTTTCTTCTCAGGATCTTTTGTGCTGCTATATTTCTGCTAAAGA 628
Db 877 TTGCTGCTGAGGAGGCAATTTCTTTTCCGGTCTTTTGGCTGCTGATATTTCTGCTCAAGA 936
Qy 629 AGAGAGGCTTATGCGAGGACTCAGTTTTCCTCAATGAACCTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 996
Qy 689 TTCACCTGCTGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 748
Db 997 TACACTGTGATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1056
Qy 749 GGGTCAGGAGATCATTTGTTGATGCTGCTCAAAATTCAGCAGGAGGTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATAATTAATCAATGCTGTTCCGATAGAACAGGAGTTCCTCACTGAGCCT 1116
Qy 809 TGCCAGTTGGCCTCATTTGGAATGAATTTGATGAACAGTACATTTGATTTGTAG 868
Db 1117 TGCTGTGAGCTCATTTGGATGAATTCGACTCTAATGAAGCAATACATTTGATTTGTGG 1176
Qy 869 CTGACAGATTACTTTGTGGAATTTGGATTCTCAAGGTTTTCAGGCAGAAATTCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACCTGGGTTTTCAGGAGGTTTTCAGAGTAGAACCCATTG 1236
Qy 929 ATTTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTTGTGAAACAGAGTTTCAG 988
```

Db 1237 ACTTATCGAGATATTTTCACTGGAAGGAAAGACTTAACCTTCTTTGAGAGAGAGTAGGCG 1296
 QY 989 AGTATCAGCGTTTTCAGTATGCGAGAAACACAGATTAACGCTTTCACCTTGGATGCGAG 1048
 Db 1297 AGTATCAGAGGATGGAGTGATGTCGAAGTCAACAGAGAAATCTTTTACCTTGGATGCTG 1356
 QY 1049 ATTT 1052
 Db 1357 ACTT 1360

RESULT 7

US-09-962-677-9
 ; Sequence 9, Application US/09962677
 ; Patent No. 6759200
 ; GENERAL INFORMATION:
 ; APPLICANT: Stanton, Jr., Vincent P.
 ; TITLE OF INVENTION: THYMIDINE PHOSPHORYLASE GENE SEQUENCE
 ; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING
 ; TITLE OF INVENTION: THE TREATMENT OF DISEASE
 ; FILE REFERENCE: 11926-015003
 ; CURRENT APPLICATION NUMBER: US/09/962,677
 ; CURRENT FILING DATE: 2001-09-24
 ; PRIOR APPLICATION NUMBER: 09/658,659
 ; PRIOR FILING DATE: 2000-09-08
 ; PRIOR APPLICATION NUMBER: 09/596,033
 ; PRIOR FILING DATE: 2000-06-15
 ; PRIOR APPLICATION NUMBER: 09/357,743
 ; PRIOR FILING DATE: 1999-07-20
 ; PRIOR APPLICATION NUMBER: 09/357,024
 ; PRIOR FILING DATE: 1999-07-19
 ; PRIOR APPLICATION NUMBER: 60/093,484
 ; PRIOR FILING DATE: 1998-07-20
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 9
 ; LENGTH: 2500
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 128..1464
 ; OTHER INFORMATION: n = g or a
 ; NAME/KEY: misc_feature
 ; LOCATION: 189
 ; OTHER INFORMATION: n = t or g
 ; NAME/KEY: misc_feature
 ; LOCATION: 524
 ; OTHER INFORMATION: n = c or g
 ; NAME/KEY: misc_feature
 ; LOCATION: 1399
 ; OTHER INFORMATION: n = t or a
 ; NAME/KEY: misc feature
 ; LOCATION: 1636..1738, 2259
 ; OTHER INFORMATION: n = c or t
 ; US-09-962-677-9

Query Match 55.1%; Score 580.6; DB 4; Length 2500;
 Best Local Similarity 75.1%; Pred. No. 2.2e-165;
 Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
 QY 89 ATGAAGAGCCATCTCTAAGAAAGAGTTCTCGCGGTTTGTGTCATCTTTTCCAATCCAGTACC 148
 Db 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTGTCATCTTCCCATCGAGTACC 456
 QY 149 CTGATATTTGGAATATGATAACAGGCACAGCTTCTCTCGACAGCAGAGAGGTTG 208
 Db 457 ATGATATCTGGAGATGTATAGAAGGCGAGAGGCTTCTCTTTGGACCGCGGAGGAGTTG 516
 QY 209 ACTATCAAGGATCTCCCTCACTGGAACAAGCTTAAGCAGATGAGAAAGTACTTCATCT 268
 Db 517 ACCTCTCNAAGACATTGACACTGGGAATCTCTGNAACCCGAGGAGAGATATTTATAT 576

QY 269 CTCACATCTTAGCCTTTTTCAGCCAGTGAGGANTCTTAATGAAATTTTGGTGAGC 328
 Db 577 CCATGTTCTTGGCTTTCTTTTCAGCAAGCGATGGCATATGTAATGAAATCTTGGTGAGC 636
 QY 329 GCTTTAGTCAGGAGTGCAGGCTTCAGAGGCTCGCTGTTTCTATGCTTTTCAAAATCTCA 388
 Db 637 GATTTAGCCAGAGTTTCAGATTACAGAAGCCGCTGTTTCTATGCTTTCCAATTTGCCA 696
 QY 389 TCGAGAATGTTCACTCAGAGATGTACAGTTTGTGTGATAGACACTTATCATCAGAGATCCCA 448
 Db 697 TGGAAAAACATACATTCTGAAATGTATAGTCTTCTTATTCACACTTACATAAAAGATCCCA 756
 QY 449 AGAAAAGGGAATTTTATTAATGCAATTTGAACCATGCGCTTGTGTCAAGAAAGAGCAG 508
 Db 757 AAGAAAGGGAATTTCTCTTCAATGCAATTTGAACCATGCGCTTGTGTCAAGAAAGAGCAG 816
 QY 509 ATTGGGCTTTGGCATGATAGCAGATAGAAAATCTACTTTTGGGAAAGAGTGGTGGCT 568
 Db 817 ACTGGGCTTTGGCTGGATTTGGGACAAAGAGCTACCTATGCTGAACGTTGTGTAGCCT 876
 QY 569 TTGCTGCTGTAGAAGAGTTTCTTCTCAGAGATCTTTTCTCAGATCTTATTTCTGCTAAAGA 628
 Db 877 TTGCTGCTGTAGAAGAGTTTCTTCTCAGAGATCTTTTCTCAGATCTTATTTCTGCTCAAGA 936
 QY 629 AGAGAGGCTTATGCCAGGACTCACTTTTCCAAATCACTCATCAGCAGAGATGAGGAC 688
 Db 937 AACGAGGACTGATGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 996
 QY 689 TTTCACTGTGACTTTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 748
 Db 997 TACACTGTGATTTTGTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1056
 QY 749 GGGTCAGGAGATCATTTGTTGATGCTTCAAAATTCAGCAGGAGTTTAAACAGAGCCT 808
 Db 1057 GAGTAAGAGAAATTAATCAATGCTTTCGGATAGAACAGGAGTTTCTTCACTGAGGCT 1116
 QY 809 TGCCAGTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 868
 Db 1117 TGCCAGTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1176
 QY 869 CTGACAGATTACTTTGGAACCTTGGATTTCTCAAGGTTTTCAGGACAGAAATCTTTTGG 928
 Db 1177 CAGACAGACTTATGCTGGAACCTTGGGTTTTCAGAGGTTTTCAGAGTACAGAACCCATTG 1236
 QY 929 ATTTTATGGAACCAATTTCTTTTAGAGGAAACAAATTTCTTTTGAACACGAGTTTCAG 988
 Db 1237 ACTTTATGGAATAATTTCACTGGAAGGAAAGACTTAATTTCTTTTGAAGAGAGTAGGCG 1296
 QY 989 AGTATCAGCGTTTTCAGATTATGGCAGAAACACAGATTAACGCTTTCACCTTGGATGCG 1048
 Db 1297 AGTATCAGAGGATGGGAGTGTCAAGTCCAACAGAGAAATTTCTTTTACCTTGGATGCTG 1356
 QY 1049 ATTT 1052
 Db 1357 ACTT 1360

RESULT 8

US-09-023-655-1370
 ; Sequence 1370, Application US/09023655
 ; Patent No. 6607879
 ; GENERAL INFORMATION:
 ; APPLICANT: Cocks, Benjamin G.
 ; APPLICANT: Susan G. Stuart
 ; APPLICANT: Jeffrey J. Seilhamer
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
 ; TITLE OF INVENTION: EXPRESSION
 ; NUMBER OF SEQUENCES: 1508
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO

STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,655
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1370:
SEQUENCE CHARACTERISTICS:
LENGTH: 2500 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g36154
US-09-023-655-1370

Query Match 55.1%; Score 580; DB 4; Length 2500;
Best Local Similarity 75.1%; Pred. No. 3.4e-165;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCCTCTAAGAAAGAGGTTCTCGCGGTTTGATCTTTTCAATCCAGTACC 148
DB 397 AGGATGAGCGGCTGTGAGAGAAAACCCCGCGCTTTGTCTATCTTCCCATCGAGTACC 456
QY 149 CTGATATTGGAAAATGTTATATAACAGCAAGGCTTCTTCTTGACAGCAGAGAGGTTG 208
DB 457 ATGATATCTGGCAGATGTTATGAAGGAGGAGGCTTCTTCTTGACCGCGAGGAGGTTG 516
QY 209 ACTTATCAAGGATCTCCCTCACTGGAACAAAGCTTAAAGCAGATGAGAACTTCTCATCT 268
DB 517 ACCTCTCCAGGACATTCAGCACTGGGAATCCTGAAACCCGAGGAGATATTTATAT 576
QY 269 CTCACATCTTAGCCCTTTTTCAGCGAGTGATGGAATGTAATGAAAATTTGGTGAGC 328
DB 577 CCCATGTTCTGGCTTTCTTTCAGCAAGCGATGGCATAGTAAATGAAAATTTGGTGAGC 636
QY 329 GCTTTAGTCAGAGGTGCGAGGTTCCAGAGGCTCGCTGTTCTATGGCTTTCAAACTCA 388
DB 637 GATTTAGCAAGAAGTTCAGATTAAGAAGCCCGCTGTTCTATGGCTTTCCAAATTTGCCA 696
QY 389 TCGAAGATGTTTCACCTCAGAGATGTTACAGTTTCTGTATGACACTTACATCAGAGATCCCA 448
DB 697 TGGAAAACATACATCTGAAATGTTATGTTCTTATTGACACTTACATAAAGATCCCA 756
QY 449 AGAAAAGGGAATTTTATTAATGCAATGAAACCATGCCCCTATGTTAAGAAAAGCAAG 508
DB 757 AAGAAAAGGGAATTTCTCTTCAATGCCATTTGAAACGATGCTGTTGTCAAGAAAGGCGAG 816
QY 509 ATTTGGCCTTGGATAGCAGATAGAAAATCTACTTTTGGGGAAGAGGTTGGGCT 568
DB 817 ACTGGGCTTGGCTTGGATTTGGGGAAGAGGCTACCTATGTTGAAACGTTGTAGGCT 876
QY 569 TTGCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTCTGCTATATTTCTGCTGCTAAAGA 628
DB 877 TTGCTGAGTGAAGGCAATTTCTTTTCGGTTCTTTTTCGGTCTGATATTTCTGGCTCAAGA 936

QY 629 AGAGAGGCTTTATGTCAGAGGACTCACTTTTCCAAATGAACATCATCAGCAGAGATGAAGGAC 688
DB 937 AACGAGACTGATGCTGGGCTCACAATTTCTAATGAACCTTATTAGCAGAGATGAGGGTT 996
QY 689 TTCACCTGTGACTTTGCTTTGCTGATGTTTCAATATCTTAGTAAATAAGCCTTTCAGAAAGAA 748
DB 997 TACACTGTGATTTTGTCTGCTGATGTTTCAACACCTGTTACACAAACCATCGGAGGAGA 1056
QY 749 GGGTCAGGAGATCATTTGTTGATGCTCTCAAAATTTAGCAGGAGGTTTTTAACAGAGCCT 808
DB 1057 GAGTAAGAGAAATAAATTTATCAATGCTGTTCGGATAGAACAGGAGTTCTCTCACTGAGGCT 1116
QY 809 TGGCAGTTGGCCTCATTTGGAATGAATTTGCAATTTTGTAGTAAACAGTACATTCAGTTTGTAG 868
DB 1117 TGCCTGTGAAGCTCAITGGGATGAATTTGCACTCTAATGAAGCAATACATTTGAGTTTGG 1176
QY 869 CTGACAGATTACTTTGTGGAATCTTGGATTTCTCAAAAGGTTTTTCAGGCGCAGAAAATCCTTTG 928
DB 1177 CAGACAGACTTATGCTGGAATCTGGTTTTCAGCAAGTTTTCAGAGTAGAAGAACCATTTG 1236
QY 929 ATTTTATGAAAACATTTCTTTTGAAGAAAACAAAATTTCTTTGAGAAAACGAGTTTTCAG 988
DB 1237 ACTTTATGGAGAATAATTTCACTGGAAGAAAAGACTAACTTCTTTGAGAAGAGAGTAGGCG 1296
QY 989 AGTATCAGCGTTTTCAGATTATGGCAGAAACCAAGATACGTTTCTTACCTTTGGATGCTG 1048
DB 1297 AGTATCAGAGGATGGGAGTGTATGCAAGTCCACAGAGAAATTTCTTTTACCTTTGGATGCTG 1356
QY 1049 ATTT 1052
DB 1357 ACTT 1360

RESULT 9

US-09-949-016-145
; Sequence 145, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 145
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-145

Query Match 55.1%; Score 580; DB 4; Length 2500;
Best Local Similarity 75.1%; Pred. No. 3.4e-165;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCCTCTTAAAGAAAGAGTTCTCGCGGTTTGTCTATCTTCCAAATCCAGTACC 148
DB 397 AGGATGAGCGGCTGTGAGAGAAAACCCCGCGCTTTGTCTATCTTCCCATCGAGTACC 456
QY 149 CTGATATTGGAAAATGTTATATAACAGCAAGGCTTCTTCTTGACAGCAGAGAGGTTG 208
DB 457 ATGATATCTGGCAGATGTTATGAAGGAGGAGGCTTCTTCTTGACCGCGAGGAGGTTG 516
QY 209 ACTTATCAAGGATCTCCCTCACTGGAACAAAGCTTAAAGCAGATGAGAACTTCTCATCT 268

```

Db 517 ACCTCTCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGATATTTTATAT 576
QY 269 CTCACATCTTAGCCTTTTTCAGCCAGTGATGGAATGTAAATGAAATTTTGGTGAGC 328
Db 577 CCCATGTTCTGGCTTTTTCAGCAAGCGATGGCATAGTAATGAAATCTTGGTGAGC 636
QY 329 GCTTTAGTCAGGAGTGCGAGGTTTCAGAGGCTCGCTGTTTCTATGCTTTCCTCAATTTCTCA 388
Db 637 GATTTAGCAAGAAGTTCAGATTCAGATTCAGAGCCGCTGTTTCTATGGCTTCCTCAATTTGCA 696
QY 389 TCGAATGTTTCTACACAGATGTCAGTTTCTGATGACACTTACATCAGAGATCCCA 448
Db 697 TCGAATGTTTCTACACAGATGTCAGTTTCTGATGACACTTACATCAGAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTAATGCAATTTGAACCAATGCCCTATGTTAAGAAAGGAGCAG 508
Db 757 AAGAAAGGGAATTTTCTCTCAATGCTTGAACCAATGCCCTATGTTAAGAAAGGAGCAG 816
QY 509 ATTGGGCTTGGATGGATAGCAGATAGAAATCTACTTTTGGGAAAGAGTGGTGCCCT 568
Db 817 ACTGGGCTTGGCTGGATGGGGAAGAGGCTACCTATGCTGAACTGATGCTGATGCTG 876
QY 569 TTGCTGCTGTAAGAGGATTTTCTTCTCAGGATCTTTTCTGCTGCTATATTTCTGCTTAAAGA 628
Db 877 TTGCTGCAAGTGAAGGCAATTTTCTTTTTCGGGTTCTTTTTCGGTCTGATATTTCTGGCTCAAGA 936
QY 629 AGAGAGGCTTATGTCAGGACTCAGTTTTCCTCAATGAACTATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCTGGCTTCAATTTTCTTAATGAACTTATAGCAGAGATGAGGGTT 996
QY 689 TTCACGTGACCTTTCCTGCTGATGTTTCCAAATCTAGTAATTAAGCCCTCAGAGAA 748
Db 997 TACACTGTGATTTTCTGCTGATGTTTCCAAATCTAGTAATTAAGCCCTCAGAGAA 1056
QY 749 GGGTCAGGAGATCATTTGTTGATGCTGCTCAAAATTTGAGCAGAGGAGTTTAAACAGAGCCCT 808
Db 1057 GAGTAAAGAAATTAATTAATGCTGTTGCGATAGAAAGAGGAGTTTCTCCTCAGTGGCCT 1116
QY 809 TGCCAGTTGGCTCATTTGAAATGAAATTTGATGAAACAGTACATTTGATGTTTGTAG 868
Db 1117 TGCCTGTGAAGCTCATTTGGATGAAATTTGCACTCTAATGAAGCAATACATTTGATTTGTGG 1176
QY 869 CTGACAGATTTCTTGGAACTTTGGATTTCTCAAGGTTTTCAGGCAAGAAATTCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACTGGGTTTGTAGCAAGGTTTTCAGAGTGAAGACCCCATTTG 1236
QY 929 ATTTTATGAAACATTTCTTTTGAAGGAAACAAATTTCTTTTACCTTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

```

```

RESULT 10
US-09-949-016-2025
; Sequence 2025, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

```

```

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2025
; LENGTH: 2479
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-2025

Query Match      54.9%; Score 578.4; DB 4; Length 2479;
Best Local Similarity 75.0%; Pred. No. 1e-164;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGCCGGTTTGTTCATCTTTCCCAATCCAGTACC 148
Db 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTCTATCTTCCCATCGAGTACC 456
QY 149 CTGATATTTTGGAAATGTATAACAGGCACAGCTTCTCTCGACAGCAGAGAGGTTG 208
Db 457 ATGATATCTGGCAGATGTATAAGAGGCAGAGGCTTCTCTTTGGACCGCGAGAGGTTG 516
QY 209 ACTTATCAAGAGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAACTACTTCACT 268
Db 517 ACTCTCCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
QY 269 CTCACATCTTAGCCTTTTTCAGCCAGTGTAGGAATTTGTAATGAAATTTGGTGGAGC 328
Db 577 CCAATGTTCTGGCTTTTTCAGCAAGGATGGCATAGTAATGAAATTTGGTGGAGC 636
QY 329 GCTTTAGTCAGAGGTCAGGTTCCAGAGGCTCGCTGTTCTATGGCTTTCAAAATTTCTCA 388
Db 637 GATTTAGCCAAGAAAGTTTCAGATTTACAGAAAGCCGCTGTTCTATGGCTTCCAAATTTGCCA 696
QY 389 TCGAATGTTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCA 448
Db 697 TCGAATGTTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTTAATGCAATTTGAAACCAATGCTTATGTTAAGAAAGGAGCAG 508
Db 757 AAGAAAGGGAATTTTCTTCTCAATGCCATTTGAAACGATGCTTGTGTCAAGAAAGGAGCAG 816
QY 509 ATTGGGCTTGGATGGATAGCAGATAGAAATCTACTTTTGGGAAAGAGTGGTGGCT 568
Db 817 ACTGGGCTTGGCTGGATGGGGAACAAAGAGGCTACTTATGTTGAACGTTGTGTAGCCT 876
QY 569 TTGCTGCTGTAGAAGAGTTTCTTCTCAGGATCTTTTGTCTGATATATTTCTGGCTTAAAGA 628
Db 877 TTGCTGCTGTAGAAGAGTTTCTTCTCAGGATCTTTTGTCTGATATATTTCTGGCTTAAAGA 936
QY 629 AGAGAGGCTTATGTCAGGACTCAGTTTTCCTCAATGAACTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCTGGCTTCAATTTTCTAATGAACTTATTAGCAGAGATGAGGGTT 996
QY 689 TTCACGTGACCTTTCCTGCTGATGTTTCCAAATCTAGTAATTAAGCCCTCAGAGAA 748
Db 997 TACACTGTGATTTTCTGCTGATGTTTCCAAATCTAGTAATTAAGCCCTCAGAGAA 1056
QY 749 GGGTCAGGAGATCATTTGTTGATGCTGCTCAAAATTTGAGCAGAGGAGTTTAAACAGAGCCCT 808
Db 1057 GAGTAAAGAAATTAATTAATGCTGTTGCGATAGAAAGAGGAGTTTCTCCTCAGTGGCCT 1116
QY 809 TGCCAGTTGGCTCATTTGAAATGAAATTTGATGAAACAGTACATTTGATGTTTGTAG 868
Db 1117 TGCCTGTGAAGCTCATTTGGATGAAATTTGCACTCTAATGAAGCAATACATTTGATTTGTGG 1176
QY 869 CTGACAGATTTCTTGGAACTTTGGATTTCTCAAGGTTTTCAGGCAAGAAATTCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACTGGGTTTGTAGCAAGGTTTTCAGAGTGAAGACCCCATTTG 1236
QY 929 ATTTTATGAAACATTTCTTTTGAAGGAAACAAATTTCTTTTACCTTTGGATGCTG 988

```

Db 1237 ACTTTATGGAGATATTTTCACTGGAAGGAAAGACTAACTTTCTTTTGAGAAAGAGTAGGCG 1296
QY 989 AGTATCAGCGTTTTCAGTTATGCGAGAAACACAGATAACGCTTCACCTTCGATGCGAG 1048
Db 1297 AGTATCAGAGATGGAGATGATGCAAGTCAACAGAGAAATCTTTTACCTTGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 11
US-08-905-223-125
; Sequence 125, Application US/08905223
; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobb, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 125:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 481 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: CDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Brain
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 41..343
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 4.4
; OTHER INFORMATION: seq ISHILAFFAASDG/IV
US-08-905-223-125

Query Match 41.7%; Score 439.4; DB 3; Length 481;
Best Local Similarity 99.8%; Pred. No. 5.9e-123;
Matches 440; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGGCGACCCGGAAGCGGAGCGGGCTGGATCAGATCAGATCATTTCA 60
Db 41 ATGGGCGACCCGGAAGCGGAGCGGGCTGGATCAGATCAGATCATTTCA 100
QY 61 GACACCAACGAAGTGAATAAAGTCAAAATGAAGGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 101 GACACCAACGAAGTGAATAAAGTCAAAATGAAGGCCACTCTTAAGAAAGAGTTCTCGC 160

QY 121 CGTTTTCTCATCTTTTCAATCCAGTACCCTGATATTTGGAAAAATGTATAAACAGGCACAG 180
Db 161 CGTTTTCTCATCTTTTCAATCCAGTACCCTGATATTTGGAAAAATGTATAAACAGGCACAG 220
QY 181 GCTTCTTCTGGACAGCAGAGAGGTTGACTTATCAAGAGGATCTCCCTCAGTGGAAACAAG 240
Db 221 GCTTCTTCTGGACAGCAGAGAGGTTGACTTATCAAGAGGATCTCCCTCAGTGGAAACAAG 280
QY 241 CTTAAAGCAGATGAGAAAGTACTTCTCATCTCTCACATCTTTAGGCTTTTTCGAGCCAGTGAT 300
Db 281 CTTAAAGCAGATGAGAAAGTACTTCTCATCTCTCACATCTTTAGGCTTTTTCGAGCCAGTGAT 340
QY 301 GGAATTCTTAATGAAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGCT 360
Db 341 GGAATTCTTAATGAAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGCT 400
QY 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db 401 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 460
QY 421 CTCATAGACACTTACATCAGA 441
Db 461 CTCATAGACACTTACATCAGA 481

RESULT 12
US-08-307-499-1/c
; Sequence 1, Application US/08307499
; Patent No. 5651972
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,499
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

```

; LENGTH: 14176 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3852..4226
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4585..4887
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5131..5310
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5760..5912
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6786..7130
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 10148..10513
; US-08-307-499-1

Query Match      40.8%; Score 429.8; DB 1; Length 14176;
Best Local Similarity 66.2%; Pred. No. 3.3e-119; Indels 5; Gaps 2;
Matches 651; Conservative 0; Mismatches 327;

QY 71 AAAGTGAATAAAGTCAATGAAGAGCCACTCTAAGAAAGAGTCTCGCCGGTTGTCA 130
Db 10841 AAAATGATTTTATACAGTAAGAGCCTATCTTCAAGAGCTGATCTAGTTCGTTA 10782

QY 131 TCTTCCAAATCCAGTACCCTGATATTTGGAAAATGATATAAAGGACAGGCTTCTTCT 190
Db 10781 TTTTCCCTATTAAAGTATCATGATATCTGAAAAATGTATAAACAATCAGTGGCAAGTTT 10722

QY 191 GGACGACGACGAGGTGACTTATCAAGAGGATCTCCCTCACTGGACACAGCTTAAAGCAG 250
Db 10721 GGACCGTTGAAGAAGTAGATTTATCAAAAGATTTAGATGATTTGGGATAAATTAACATAAG 10662

QY 251 ATGAGAAGTACTTCTCTCATCTTAGCCCTTTTTCGACGACGATGATGAAATGTAA 310
Db 10661 ACGAAAATACCTTTATAAACAATATAGCAATTTTTCATCTAGTATGATGTATGTAA 10602

QY 311 ATGAAAAATTTGGTGAGCCCTTTAGTCAGAGGTCGAGTTCCAGAGGCTCGCTGTTCT 370
Db 10601 ATGAAAATTTAGCGGAAAAGATTTTATGTGATGTACAGTGTTCAGAGGACGATGTTCT 10542

QY 371 ATGGCTTCAAAATCTCATCGAGATGTTTCACTCAGAGATGATACAGTTTCTGATAGACA 430
Db 10541 ATGGATTTCAAAATAGCTATGGAATAATTTTATTCAGAAATGATATGATTTTAAATAGATA 10482

QY 431 CTTACATCAGAGATCCCAAGAAAAGGGAATTTTAAATGCAATTTGAACCATGCCCT 490
Db 10481 CATATGTAAGAGATAATATAGAAAAATGCAATTTTAAACCGTATAGAAACAATGGAT 10422

QY 491 ATGTTAAGAAAAAGCAGATCGGCTTGGCGATGATAGCAGATAGAAAAATCTACTTTTG 550
Db 10421 GCGTAAAAAAGAAAGCTGATTGGCCAGAAAAATGGATATC---TAGCAACAGGATATATG 10365

QY 551 GGGAAAGAGTGTGGCTTGTCTGTGTAGAAAGGATTTCTTCTCAGGATCTTTTGTG 610
Db 10364 GAGAAAGAGTAGTAGCATTTGAGCTGTGGAGGGAATATCTTTTCTGTTCAATTTGCTG 10305

QY 611 CTATATCTGGCTAAAGAGAGAGGCTTTATCCAGGACTCACTTTTCCAAATGAACCTCA 670
Db 10304 CTATATTTTGGATAAAAAACAGAGGATGTAUCCCGGATTAACAATTTCTAATGAACCTAA 10245

QY 671 TCACGAGAGATGAAGGACTTCACTGATGCTTGTGCTGCTGATGTTTCCAAATCTTAGTAA 730
Db 10244 TAAGTAGAGACCAAGGCTTTACATTTGTGATTTTGGCGTGTAAATGTTTAAACATTTATTAC 10185

```

```

QY 731 ATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAAAATTTGAGCAGG 790
Db 10184 ATCCACCATCTAAGGAAGTTATAACGTCGATATCATTTGATGCGGTAAATATAGAAAAGG 10125

QY 791 AGTTTTTAAACAGAAGCCTTGGCCAGTTGGCCTCATTCGAAATGAATTCATTTTATGAAAC 850
Db 10124 AGTTTTTGACAGTTGCTATTCGGTGGATCTTATAGGTATGAATTTGTTTAAATGTCCTC 10065

QY 851 AGTACATTCAGTTTGTAGCTGACAGATTACTTGTGAACTTGTGATCTCAAGGTTTTC 910
Db 10064 AGTATATAGAATTCGTCGAGATAGATTATTAACAGAGTTAGGTTGTGAAAAG--TCTCA 10007

QY 911 AGCGAGAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGGAAAAACAAATTTCT 970
Db 10006 ATGTATATATCTTTTGTAGCTTTTATGAGTATATATCACTAGAGGTTAGACTAATTTT 9947

QY 971 TTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGCTTATGCGAGAAACACAGATAACG 1030
Db 9946 TCGAAGCAGAGTTAGTGAATATCAAAAGATGGGGTGTTTACAAATAAAGAGAGAATA 9887

QY 1031 TCTTCACCTTGGATCGAGATTTT 1053
Db 9886 TATTTTCTACGGATATAGATTTT 9864

RESULT 13
US-08-307-499-14
; Sequence 14, Application US/08307499
; Patent No. 5651972
; GENERAL INFORMATION:
; APPLICANT: Meyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,499
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14176 base pairs

```

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS
LOCATION: 138..1460

FEATURE:

NAME/KEY: CDS
LOCATION: 2456..2659

FEATURE:

NAME/KEY: CDS
LOCATION: 2809..3030

FEATURE:

NAME/KEY: CDS
LOCATION: 3070..3330

FEATURE:

NAME/KEY: CDS
LOCATION: 3356..4180

FEATURE:

NAME/KEY: CDS
LOCATION: 4392..5894

FEATURE:

NAME/KEY: CDS
LOCATION: 6171..6398

FEATURE:

NAME/KEY: CDS
LOCATION: 6447..6875

FEATURE:

NAME/KEY: CDS
LOCATION: 6928..7431

FEATURE:

NAME/KEY: CDS
LOCATION: 7454..7858

FEATURE:

NAME/KEY: CDS
LOCATION: 7895..8155

FEATURE:

NAME/KEY: CDS
LOCATION: 8215..8682

FEATURE:

NAME/KEY: CDS
LOCATION: 8715..9539

FEATURE:

NAME/KEY: CDS
LOCATION: 9562..10272

FEATURE:

NAME/KEY: CDS
LOCATION: 10316..11908

FEATURE:

NAME/KEY: CDS
LOCATION: 11971..12780

FEATURE:

NAME/KEY: CDS
LOCATION: 12829..13107

FEATURE:

NAME/KEY: CDS
LOCATION: 13149..14171

FEATURE:

NAME/KEY: CDS
LOCATION: 13149..14171

FEATURE:

NAME/KEY: CDS
LOCATION: 13149..14171

FEATURE:

NAME/KEY: CDS
LOCATION: 13149..14171

FEATURE:

NAME/KEY: CDS
LOCATION: 13149..14171

FEATURE:

NAME/KEY: CDS
LOCATION: 13149..14171

Query Match 40.8%; Score 429.8; DB 1; Length 14176;

Best Local Similarity 66.2%; Pred. No. 3.3e-119;

Matches 651; Conservative 0; Mismatches 327; Indels 5; Gaps 2;

71 AAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAGAGTTCTCCCGGTTGTCA 130

3336 AAATGATTTTATACGAATGGAGCCCTATTCTCAAGAGTCTGATTTCTAGTTCGTTA 3395

131 TCTTCCCAATCCAGTACCTGATATTGGAAATGTATAACAGGACAGGCTTCCTCT 190

3396 TTTTCCCTATTAGTATCATGATATCTCGAAATGTATAACAATCAGTGGCAAGTTT 3455

191 GGACAGCAGAAGAGTTGATCTTCAAGGATCTCCCTCACTGGAAACAAGCTTAAGCAG 250

3456 GGACCGTTGAAGAGTAGATTTCATCAAAAGATTAGATGTTGGATAAATAACTAAAG 3515

251 ATGAGAAGTACTTCACTCTCAATCTTACATCTTTTGGCAGCCAGTATGGAATGTGAA 310

3516 ACGAAAAATACTTTTATAAAAAATATAGTATCTAGTATCTTTTGGATGATGATGTTG 3575

311 ATGAAAATTTGGTGGAGCGCTTTAGTTCAGAGGTGAGGTTCAGAGGCTCGCTGTTTCT 370

3576 ATGAGAATTTTGGGAAAGATTATGATGATGATGATGATGATGATGATGATGATGAT 3635

371 ATGGCTTTTCAAAATTTCTCATCGAGATGTTCACTCAGAGATGATCAGTTCGTGATGACA 430

3636 ATGGATTTCAATAGCTATGGAATAATTCATTCAGAAATGATAGTATTATTAATAGATA 3695

431 CTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATGAAACCAATGCCCT 490

3696 CATATGTAAGAGATAATATAGAAAAAATGCAATTTATTAACGCTATAGAAAAATGGAAT 3755

491 ATGTTAAGAAAAAGCAGATTTGGCCTTGGCCTTGGATGATGATGATGATGATGATGATG 550

3756 GCGTAAAAAAGAAAGCTGATTGGGCCAGAAAAATGGATATC--TAGCAACAAGGTATATG 3812

551 GGGAAAGAGTGGTGGCTTTGCTGCTAGAGAGGATTTTCTCTCAGGATCTTTTGGCTG 610

3813 GAGAAAGAGTAGTAGCATTTGCGCTGTGAGGGAATATTCCTTTCTGTTCAATTTGCTG 3872

611 CTATATTTCTGGCTAAAGAAAGAGAGGCTTTATGATGATGATGATGATGATGATGATGAT 670

3873 CTATATTTTGGATAAAAAAGCAGAGATTGATGATGATGATGATGATGATGATGATGAT 3932

671 TCAGCAGAGATGAAGGACTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 730

3933 TAAAGTAGAGACGAAGGTTTACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3992

731 ATAAGCTTCAGAGAAAGGCTCAGGAGATCATTTGATGCTGCTGCTGCTGCTGCTGCTGCT 790

3993 ATCCACCATCTAAGGAAGTTTAACTGATGATGATGATGATGATGATGATGATGATGATG 4052

791 AGTTTTTAAACAGAAAGCTTTCCAGTTGGCTTCATTTGGAATGAAATGATGATGATGATG 850

4053 AGTTTTTGACAGTTGCTATTTCCGTTGGATCTTATAGGTATGATGATGATGATGATGAT 4112

851 AGTACATGAGTTTGTAGCTGACAGATTAATCTGTTGGAATCTTGAAGTTCTCAAGGTTTTC 910

4113 AGTATATAGAAATTCGTCGACAGATAGATTATTAACAGAGTTAGGTTGTGAAAAAG--TCTCA 4170

911 AGGCAGAAATCCTTTTGTATTTTATGAAAAACATTTCTTTTAGAGGAAAAACAATTTCT 970

4171 ATGTATATATCTCTTTTAGCTTTTATGAGTATATATACAGAGGTAAGACTAATTTT 4230

971 TTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTCATGCGAGAAACACAGATACG 1030

4231 TCGAACGACGAGTTAGTGAATATCAAAAGATGGGGTGTTTTACAAATTAAGAGAGAATA 4290

1031 TCTTCACCTTGGATGAGATTTT 1053

4291 TATTTTCTACGATATAGATTTT 4313

RESULT 14

US-09-299-268-1/c

; Sequence 1, Application US/09299268

; Patent No. 6217882

; GENERAL INFORMATION:

; APPLICANT: Moyer, Richard W.

; APPLICANT: Vi uela, Eladio

; APPLICANT: Gibbs, E.P.J.

; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a

; TITLE OF INVENTION: Live Vaccine Vector

; NUMBER OF SEQUENCES: 60

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: David R. Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.
ZIP: 32606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,268
FILING DATE: US/09/299,268

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/901,127
FILING DATE:
APPLICATION NUMBER: US 07/908,241
FILING DATE: 1-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,630
FILING DATE: 29-JUN-1992

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/342,212
FILING DATE: 21-APR-1992
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sallwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 14176 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 3852..4226

FEATURE:
NAME/KEY: CDS
LOCATION: 4585..4887

FEATURE:
NAME/KEY: CDS
LOCATION: 5131..5310

FEATURE:
NAME/KEY: CDS
LOCATION: 5760..5912

FEATURE:
NAME/KEY: CDS
LOCATION: 6786..7130

FEATURE:
NAME/KEY: CDS
LOCATION: 10148..10513

Query Match 40.8%; Score 429.8; DB 3; Length 14176;
Best Local Similarity 66.2%; Pred. No. 3.3e-119;
Matches 651; Conservative 0; Mismatches 327; Indels 5; Gaps 2;

71 AAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTCTCGCCGGTTGTCA 130
10841 AAATGATTTTATACAGCAATGGAGCCCTATTCTTCAAGAGTCTGATTTAGTTCGTTA 10782
131 TCTTTCCATCCAGTACCCCTGATATTGGAAATGTATATAACAGGCACAGGCTTCCTTCT 190
10781 TTTTCCCTATTAAAGTATCATGATATCTGAAAAATGTATAACAATCAGTGGCAAGTTT 10722

191 GGACAGCAGAAAGAGTTGACTTATCAAAAGATCTCCCTCACTGGAAACAAGCTTAAGCAG 250
10721 GGACCGTTGAAGAGTAGATTATCAAAAGATTTAGATGATGGATATAATTAACATAAG 10662
251 ATGAGAAAGTACTTCTCTCTACATCTTAGCCCTTTTGGAGCCAGTGTAGGAATGTGAA 310
10661 ACCAAAATACTTTATATAAAACATATACTAGCATTTTGGATCTAGTGTAGTGTGAA 10602
311 ATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCTCGCTGTTCT 370
10601 ATGAGAAATTTAGCGGAAAGATTTTATGTGATGTACAGTGTTCAGAGGCACGATGTTCT 10542
371 ATGGCTTTCAAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTGTGTATAGCA 430
10541 ATGGATTTCAAATAGCTATGGAATAATTCTTCAGAAATGTATAGTTTATTAATAGATA 10482
431 CTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTTGAACCAATGCGCT 490
10481 CATATCTAAGAGATAATATAGAAAAAATGCAATTTATTTAAGCTTATAGAAAAAATGGAAT 10422
491 ATGTTAAGAAAAAGCAGATTTGGCCCTTGCATGCGATGGATAGCAGATAGAAAAATCTACTTTG 550
10421 GCGTAAAAAAGAAAGCTGATTTGGGCCAGAAAAATGGATATC--TAGCAACAAGGTATATG 10365
551 GGGAAAGAGTGTGGCTTTGCTGCTGTAGAAAGAGTCTTCTCTCAGGATCTTTTGTCTG 610
10364 GAGAAAGAGTAGTAGCATTTGCACTGTGAGGGAATATTCTTTCTGTTTCAATTTGCTG 10305
611 CTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGAGCTCACCTTTTCCAATGAATCA 670
10304 CTATATTTGGATAAAAAAGCAGAGATTTGATGCCGGATTAACAATTTTCTTAATGAATCA 10245
671 TCAGCAGAGATGAGGACTTCACTGTGACTTTTCTGCTGCTGCTGCTTCCATATCTAGTAA 730
10244 TAAGTAGACGACGAAGGTTTACATTTGCTGTTTAAAGTGTAAACATTTATTTAC 10185
731 ATAAAGCTTTCAGAAAGAAAGGTTCAGGAGATCATTTGTGATGCTGTCAAAATTCAGCAGG 790
10184 ATCCACCATCTAAGAAAGTTATAACGTCGATATCATTTGATCGGTTAATATAGAAAAAG 10125
791 AGTTTTTAACAGAAAGCTTTGCCAGTTGGCTCTCATTTGGAATGAATTTGATGAAAC 850
10124 AGTTTTTGACAGTTGCTATTCCGGTGGATCTTATAGGTATGAATTTGTTTAAATGCTC 10065
851 AGTACATTTGAGTTTGTAGCTGACAGATTAATTTGGAAGTCTTCAAGGTTTTC 910
10064 AGTATATAGAAATTCGTGCGAGATAGATTTAACAAGAGTTAGGTTGTGAAAAAG--TCTCA 10007
911 AGGCAGAAATCCTTTGATTTTATGAAAAATTTCTTTAGAAAGAAAAACAATTTCT 970
10006 ATGTATATATCTCTTTTAGCTTTATGGAGTATATATCCTAGAGGTAAGACTAATTTT 9947
971 TTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGTTTTCAGTTTGGCAGAAACCAAGATAACG 1030
9946 TCGAACGACGAGTTAGTGAATATCAAAAGATGGGGTGTTTTACAAATAAAGAGAGAATA 9887
1031 TCTTCACTTGGATGAGATTTT 1053
9886 TATTTTCTACGGATATAGATTTT 9864

RESULT 15
US-09-299-268-14
; Sequence 14, Application US/09299268
; Patent No. 6217882
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:

ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,268
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/901,127
FILING DATE:
APPLICATION NUMBER: US 07/908,241
FILING DATE: 1-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,630
FILING DATE: 29-JUN-1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/342,212
FILING DATE: 21-APR-1992
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 14176 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 138..1460
FEATURE:
NAME/KEY: CDS
LOCATION: 2456..2659
FEATURE:
NAME/KEY: CDS
LOCATION: 2809..3030
FEATURE:
NAME/KEY: CDS
LOCATION: 3070..3330
FEATURE:
NAME/KEY: CDS
LOCATION: 3356..4180
FEATURE:
NAME/KEY: CDS
LOCATION: 4392..5894
FEATURE:
NAME/KEY: CDS
LOCATION: 6171..6398
FEATURE:
NAME/KEY: CDS
LOCATION: 6447..6875
FEATURE:
NAME/KEY: CDS
LOCATION: 6928..7431
FEATURE:
NAME/KEY: CDS
LOCATION: 7454..7858
FEATURE:

NAME/KEY: CDS
LOCATION: 7895..8155
FEATURE:
NAME/KEY: CDS
LOCATION: 8215..8682
FEATURE:
NAME/KEY: CDS
LOCATION: 8715..9539
FEATURE:
NAME/KEY: CDS
LOCATION: 9562..10272
FEATURE:
NAME/KEY: CDS
LOCATION: 10316..11908
FEATURE:
NAME/KEY: CDS
LOCATION: 11971..12780
FEATURE:
NAME/KEY: CDS
LOCATION: 12829..13107
FEATURE:
NAME/KEY: CDS
LOCATION: 13149..14171
US-09-299-268-14
Query Match 40.8%; Score 429.8; DB 3; Length 14176;
Best Local Similarity 66.2%; Pred. No. 3.3e-119;
Matches 651; Conservative 0; Mismatches 327; Indels 5; Gaps 2;
QY 71 AAAGTGAATAAAAGTCAAAATGAAGAGCCACTCTCTCAAGAAAAGATTCTCGCGGTTTGTC 130
DB 3336 AAATGATTTTATACACGATCGAGCCTATTCTTCAAGAGTCTGATCTTAGGTCGTTA 3395
QY 131 TCTTTCCAAATCCAGTACCCCTGATATTTGGAAAATGTATAAACAGGACAGGCTTCCTTCT 190
DB 3396 TTTTCCCTATTAAAGTATCATGATATCTGAAAATGTATAAACAAATCAGTGGCAAGTTT 3455
QY 191 GGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAAGCTTAAGCAG 250
DB 3456 GGACCGTTGAAGAAAGTAGATTATCAAAAGATTAGATGATGGGATAAATTAACATAAG 3515
QY 251 ATGAGAAAGTACTTCTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGGAGTGAATGTA 310
DB 3516 ACGAAAATACTTTTATAAACATATAGTACATTTTTCAGTCTAGTATGATGTTGTA 3575
QY 311 ATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCCAGAGGCTCGCTGTTTCT 370
DB 3576 ATGAGAAATTTAGCGGAAAGATTTTATGTGGATGTACAGTGTTCAGAGGACGATGTTCT 3635
QY 371 ATGGCTTCAAATTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTGTGTATAGACA 430
DB 3636 ATGGATTTCAAATAGCTATGGAATAATTCATTTCAAGAAATGTATAGTTTATTAAGATA 3695
QY 431 CTTACATCAGAGATCCCAAGAAAAGGAATTTTATTTAATGAATGAAACCATGCCCT 490
DB 3696 CATATGTAGAGATAATATAGAAAATAATGCAATTTATTTAAGCTATAGAAAACAATGGA 3755
QY 491 ATGTTAAGAAAAAGCAGATTGGGCTTTCGATGATAGCAGATAGAAAAATCTACTTTTG 550
DB 3756 GCGTAAAAAAGAAAGCTGATTTGGCCAGAAAATGGATATC--TAGCAACAAGGTATATG 3812
QY 551 GGGAAAGAGTGGGCTTTGCTGTGTAGAAAGAGTTTCTCTCAGGATCTTTTGTGCTG 610
DB 3813 GAGAAAAGATGATGATGATTTGCACTGCTGAGGAAATTTCTTTCTGTTTCAATTTGCTG 3872
QY 611 CTATATTTCTGGCTAAAGAGAGAGGCTTATATGCCAGGACTCACATTTTCCAAATCACTCA 670
DB 3873 CTATATTTGGATAAAAAAACAGGAGATGATGCCCGGATTAACATTTTCTAATGAACATA 3932
QY 671 TCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGCTGCTTCCAAATCTAGTAA 730
DB 3933 TAAGTAGACGACGAGGTTTACATTTGTGATTTTGGGTGTTTAAATGTTTAAACATTTATAC 3992

Qy	731	ATAAGCCTTCAGAGAAAGGGTCAGGGAGATCAATTGTTGATGCTGTCAAAATTGAGCAGG	790
Db	3993	ATCCACCATCTAAGGAGTTATAACGTCGATATCATTCATGCGGTTATATAGAAAGG	4052
Qy	791	AGTTTTTAACAGAACCTTGCAGTTGGCCTCATTGGAATGAATTGCATTTTGTGATGAAC	850
Db	4053	AGTTTTTGACAGTTGCTATTCGGTGGATCTTATAGGTATGAATTGTTGTTAAATGCTC	4112
Qy	851	AGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCAGAAGTTTTTC	910
Db	4113	AGTATATAGAATTCGTCGCAGATAGATTATTAAACAGAGTTAGGTTGTGAAAAG--TCTCA	4170
Qy	911	AGGCAGAAAATCCTTTTGAATTTTATGGAACAATTTCTTTAGAAGGAAAAACAATTTCT	970
Db	4171	ATGTATATAATCCTTTTAGCTTTATGGAGTATATCACTAGAGGTAAGACTAATTTT	4230
Qy	971	TTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACCAAGATAACG	1030
Db	4231	TCGAAACGACGAGTTAGTGAATATCAAAAGATGGGGTGTTTACAATAAAGAAGAATA	4290
Qy	1031	TCTTCACCTTGGATGCAGATTTT	1053
Db	4291	TATTTTCTACGGATATAGATTTT	4313

Search completed: October 30, 2005, 03:28:15
Job time : 149 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2005, 01:51:37 ; Search time 676.5 Seconds
(without alignments)
12855.457 Million cell updates/sec

Title: US-10-698-228-12
Perfect score: 1053
Sequence: 1 atggcgaccggaagcc.....tcaccttgatgcagatttt 1053

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9784742 seqs, 4129495052 residues

Total number of hits satisfying chosen parameters: 19569484

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

4: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*

10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*

11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*

15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*

17: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*

18: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*

19: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*

20: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*

21: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*

22: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*

23: /cgn2_6/ptodata/1/pubpna/US10I_NEW_PUB.seq:*

24: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*

25: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*

26: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*

27: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*

28: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1053	100.0	1053	19	US-10-698-228-12
2	1053	100.0	1081	19	US-10-698-228-4
3	1051.4	99.8	1053	19	US-10-698-228-2
4	1051.4	99.8	4955	18	US-10-172-118-71
5	1051.4	99.8	4955	19	US-10-342-887-71

6	1051.4	99.8	4955	19	US-10-698-228-3
7	580	55.1	2482	24	US-10-220-335-514
8	580	55.1	2500	9	US-09-954-456-724
9	580	55.1	2500	9	US-09-954-456-1169
10	580	55.1	2500	9	US-09-954-456-1827
11	580	55.1	2500	19	US-10-641-643-1370
12	580	55.1	2500	21	US-10-733-878-458
13	580	55.1	2500	22	US-10-843-641A-3751
14	580	55.1	2500	22	US-10-843-641A-4196
15	580	55.1	2500	22	US-10-843-641A-4854
16	580	55.1	2500	24	US-10-756-149-713
17	578.4	54.9	1989	9	US-09-925-301-505
18	578.4	54.9	2216	16	US-10-084-817-342
19	570.4	54.2	2113	22	US-10-764-420-1636
20	569	54.0	2641	24	US-10-220-335-170
21	562.6	53.4	1328	19	US-10-403-571-75
22	545.4	51.8	977	18	US-10-264-237-790
23	514	48.8	1371	24	US-10-450-763-15278
24	463	44.0	1289	26	US-11-097-143-19181
25	447.8	42.5	186854	22	US-10-872-156-34
26	442	42.0	1146	9	US-09-822-830A-174
27	433.8	41.2	1218	16	US-10-128-714-7245
28	420.4	39.9	3945	26	US-11-097-143-19180
29	404.8	38.4	963	18	US-10-369-493-34052
30	398.6	37.9	1146	16	US-10-128-714-2245
31	394.2	37.4	1173	18	US-10-369-493-46252
32	393.6	37.4	1242	17	US-10-032-585-6537
33	388.2	36.9	1292	18	US-10-369-493-27809
34	382	36.3	1200	18	US-10-369-493-45857
35	380.6	36.1	1248	19	US-10-424-599-126843
36	372.8	35.4	1361	19	US-10-424-599-40471
37	365.8	34.7	1586	19	US-10-424-599-73024
38	363.2	34.5	957	21	US-10-653-047-43
39	363.2	34.5	1206	18	US-10-320-797-2145
40	350.6	33.3	1314	16	US-10-128-714-1245
41	350.6	33.3	1450	16	US-10-128-714-6245
42	350.6	33.3	3314	16	US-10-128-714-245
43	350.6	33.3	3450	16	US-10-128-714-5245
44	349.8	33.2	629	24	US-10-450-763-6851
45	328.2	31.2	560	9	US-09-864-761-12381

ALIGNMENTS

RESULT 1

US-10-698-228-12

Sequence 12, Application US/10698228

Publication No. US20040072253A1

GENERAL INFORMATION:

APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA

TITLE OF INVENTION: New Protein and its DNA

FILE REFERENCE: 2619W00P

CURRENT APPLICATION NUMBER: US/10/698,228

CURRENT FILING DATE: 2003-10-30

PRIOR APPLICATION NUMBER: US/10/019,733

PRIOR FILING DATE: 2001-12-28

PRIOR APPLICATION NUMBER: JP 11-181131

PRIOR FILING DATE: 1999-06-28

PRIOR APPLICATION NUMBER: JP 11-192391

PRIOR FILING DATE: 1999-07-06

PRIOR APPLICATION NUMBER: JP 2000-017770

PRIOR FILING DATE: 2000-01-21

NUMBER OF SEQ ID NOS: 14

SEQ ID NO 12

LENGTH: 1053

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION:

US-10-698-228-12

Query Match 100.0%; Score 1053; DB 19; Length 1053;

Best Local Similarity 100.0%; Pred. No. 9.3e-295;			
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGGCGGACCGGAAAGCGGAAAGCGCGGCGGTGGATCAGGATGAGAGATCATCTTCA	60
Db	1	ATGGCGGACCGGAAAGCGGAAAGCGCGGCGGTGGATCAGGATGAGAGATCATCTTCA	60
QY	61	GACACCAACGAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCG	120
Db	61	GACACCAACGAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCG	120
QY	121	CGGTTTGTCATCTTCCCAATCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG	180
Db	121	CGGTTTGTCATCTTCCCAATCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG	180
QY	181	GCTTCCTTCTGGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAAACAAG	240
Db	181	GCTTCCTTCTGGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAAACAAG	240
QY	241	CTTAAAGCAGATGAGAAGTACTCTCTCATCTTAGCCCTTTTTCAGCCAGTGAT	300
Db	241	CTTAAAGCAGATGAGAAGTACTCTCTCATCTTAGCCCTTTTTCAGCCAGTGAT	300
QY	301	GGAAATGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGT	360
Db	301	GGAAATGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGT	360
QY	361	CGCTGTTCTATGGCTTTCAAAATCTCATCGAAGTGTCTCACTAGAGATGACAGTTTG	420
Db	361	CGCTGTTCTATGGCTTTCAAAATCTCATCGAAGTGTCTCACTAGAGATGACAGTTTG	420
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA	480
Db	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA	480
QY	481	ACCATGCCCTATGTTAAAGAAAGCAGATTTGGGCTTTCGATGATAGATAGATAAAA	540
Db	481	ACCATGCCCTATGTTAAAGAAAGCAGATTTGGGCTTTCGATGATAGATAGATAAAA	540
QY	541	TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAAGAGGAGTTTCTTCTCAGGA	600
Db	541	TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAAGAGGAGTTTCTTCTCAGGA	600
QY	601	TCTTTTGTGCTATATTCTGGCTAAGAGAGAGTCTTATGCCAGGACTCATTTTTC	660
Db	601	TCTTTTGTGCTATATTCTGGCTAAGAGAGAGTCTTATGCCAGGACTCATTTTTC	660
QY	661	AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCTGATGTTCCAA	720
Db	661	AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCTGATGTTCCAA	720
QY	721	TACTTAGTAAATTAAGCCTTCAAGAAAGGGTCAAGGAGATCATTTGATGCTGCAAA	780
Db	721	TACTTAGTAAATTAAGCCTTCAAGAAAGGGTCAAGGAGATCATTTGATGCTGCAAA	780
QY	781	ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAAATGAATTGCATT	840
Db	781	ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAAATGAATTGCATT	840
QY	841	TTGATGAACACGTACATTGAGTTTGTAGCTGACAGATTACTTTGGAACTGGATTCTCA	900
Db	841	TTGATGAACACGTACATTGAGTTTGTAGCTGACAGATTACTTTGGAACTGGATTCTCA	900
QY	901	AAGGTTTTTTCAGGCGAATAATCCTTTTGATTTTATGGAAACATTTCTTTAGAAGAAA	960
Db	901	AAGGTTTTTTCAGGCGAATAATCCTTTTGATTTTATGGAAACATTTCTTTAGAAGAAA	960
QY	961	ACAAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC	1020
Db	961	ACAAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC	1020
QY	1021	ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053	

Db 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053

RESULT 2

US-10-698-228-4
; Sequence 4, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W0P
; CURRENT APPLICATION NUMBER: US/10/698,228
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-4

Query Match 100.0%; Score 1053; DB 19; Length 1081;			
Best Local Similarity 100.0%; Pred. No. 9.4e-295;			
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGGCGGACCGGAAAGCGGAAAGCGCGGCGGTGGATCAGGATGAGAGATCATCTTCA	60
Db	20	ATGGCGGACCGGAAAGCGGAAAGCGCGGCGGTGGATCAGGATGAGAGATCATCTTCA	79
QY	61	GACACCAACGAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCG	120
Db	80	GACACCAACGAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCG	139
QY	121	CGGTTTGTCATCTTCCCAATCCAGTACCTCTCATCTCTAGCCCTTTTTCAGCCAGTGAT	180
Db	140	CGGTTTGTCATCTTCCCAATCCAGTACCTCTCATCTCTAGCCCTTTTTCAGCCAGTGAT	199
QY	181	GCTTCCTTCTGGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAAG	240
Db	200	GCTTCCTTCTGGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAAG	259
QY	241	CTTAAAGCAGATGAGAAGTACTTCACTCTCATCTTAGCCCTTTTTCAGCCAGTGAT	300
Db	260	CTTAAAGCAGATGAGAAGTACTTCACTCTCATCTTAGCCCTTTTTCAGCCAGTGAT	319
QY	301	GGAAATGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGAGGCT	360
Db	320	GGAAATGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGAGGCT	379
QY	361	CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAAATGTTTCACTCAGAGATGTACAGTTG	420
Db	380	CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAAATGTTTCACTCAGAGATGTACAGTTG	439
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA	480
Db	440	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA	499
QY	481	ACCATGCCCTATGTTAAGAAAGCAGATTTGGGCTTTCGATGGATAGCAGATAGAAAA	540
Db	500	ACCATGCCCTATGTTAAGAAAGCAGATTTGGGCTTTCGATGGATAGCAGATAGAAAA	559
QY	541	TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAGAAAGAGTTTCTTCTCAGGA	600
Db	560	TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAGAAAGAGTTTCTTCTCAGGA	619

```
QY 601 TCTTTTGCTGCTATATTTCTGGCTTAAAGAGAGAGAGGCTTATGCGCAGGACTCATCTTTTCC 660
DB 620 TCTTTTGCTGCTATATTTCTGGCTTAAAGAGAGAGGCTTATGCGCAGGACTCATCTTTTCC 679
QY 661 AATGAACACTCATGACGAGAGATGAAGGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 720
DB 680 AATGAACACTCATGACGAGAGATGAAGGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 739
QY 721 TACTTAGTAAATAAGCCTTTCAGAAAGAGGCTCAGGAGATCATCTTGTGCTGTGCTCAA 780
DB 740 TACTTAGTAAATAAGCCTTTCAGAAAGAGGCTCAGGAGATCATCTTGTGATGCTGTCAA 799
QY 781 ATTGAGCAGGAGTTTAAACAGAAAGCCTTGCAGGCTTGCCTCATTTGGAATGAATTCATT 840
DB 800 ATTGAGCAGGAGTTTAAACAGAAAGCCTTGCAGGCTTGCCTCATTTGGAATGAATTCATT 859
QY 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTTCTCA 900
DB 860 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTTCTCA 919
QY 901 AAGGTTTTTCAGGAGAAATCCTTTTGATTTTATGGAACATTTCTTTAGAGGAAAA 960
DB 920 AAGGTTTTTCAGGAGAAATCCTTTTGATTTTATGGAACATTTCTTTAGAGGAAAA 979
QY 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGATATCAGCGTTTTCAGTATGCGAGAAACC 1020
DB 980 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGATATCAGCGTTTTCAGTATGCGAGAAACC 1039
QY 1021 ACAGATAAGCTCTTACCTTGGATGCGAGATTTT 1053
DB 1040 ACAGATAAGCTCTTACCTTGGATGCGAGATTTT 1072
```

RESULT 3

```
US-10-698-228-2
; Sequence 2, Application US/10698228
; Publication No. US2004007253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-2
```

```
Query Match 99.8%; Score 1051.4; DB 19; Length 1053;
Best Local Similarity 99.9%; Pred. No. 2.7e-294;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 ATGGCGGACCCGGAAAGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 1 ATGGCGGACCCGGAAAGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACCAACGAAAGTGAATAAGTCAATGAAGGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 61 GACACCAACGAAAGTGAATAAGTCAATGAAGGCCACTCTTAAGAAAGAGTTCTCGC 120
```

```
QY 121 CGGTTTGTCTATCTTTTCCATCCAGTACCTGATATTTGGAAAAATGATAAAACAGGCACAG 180
DB 121 CGGTTTGTCTATCTTTTCCATCCAGTACCTGATATTTGGAAAAATGATAAAACAGGCACAG 180
QY 181 GCTTCTCTTCTGGACAGCAGAGAGGTTGACTTATCAAAGGATCTCCCTCACTGGAAACAAG 240
DB 181 GCTTCTCTTCTGGACAGCAGAGAGGTTGACTTATCAAAGGATCTCCCTCACTGGAAACAAG 240
QY 241 CTTTAAAGCAGATCAGAGATGATCTCTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGTGTAT 300
DB 241 CTTTAAAGCAGATCAGAGATGATCTCTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGTGTAT 300
QY 301 GGAATTTGTAATCAAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
DB 301 GGAATTTGTAATCAAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
QY 361 CGCTGTTTCTATCGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
DB 361 CGCTGTTTCTATCGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTCGA 480
DB 421 CTGATAGACACTTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTCGA 480
QY 481 ACCATGCCCTTATCTTAAGAAAAAGCAGATTGGGCCCTTGGCATGGATAGCAGATAGAAAA 540
DB 481 ACCATGCCCTTATCTTAAGAAAAAGCAGATTGGGCCCTTGGCATGGATAGCAGATAGAAAA 540
QY 541 TCTACTTTTGGGAAAAAGAGTGGTGGCCTTTGCTGCTGTAGAAAGAGATTTTCTTCTCAGGA 600
DB 541 TCTACTTTTGGGAAAAAGAGTGGTGGCCTTTGCTGCTGTAGAAAGAGATTTTCTTCTCAGGA 600
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCGCAGGACTCACTTTTTCC 660
DB 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCGCAGGACTCACTTTTTCC 660
QY 661 AATGAACACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATCTTCAA 720
DB 661 AATGAACACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATCTTCAA 720
QY 721 TACTTAGTAAATAAGCCTTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
DB 721 TACTTAGTAAATAAGGCTTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
QY 781 ATTGAGCAGGAGTTTAAACAGAAAGCCTTGCAGGTTGGCCTCATTTGGAATGAATTCGATT 840
DB 781 ATTGAGCAGGAGTTTAAACAGAAAGCCTTGCAGGTTGGCCTCATTTGGAATGAATTCGATT 840
QY 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTTCTCA 900
DB 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTTCTCA 900
QY 901 AAGGTTTTTTCAGGAGAAATCCTTTTGAATTTATGAAAAACATTTCTTTAGAAAGAAAA 960
DB 901 AAGGTTTTTTCAGGAGAAATCCTTTTGAATTTATGAAAAACATTTCTTTAGAAAGAAAA 960
QY 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
DB 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
QY 1021 ACAGATAAGCTCTTACCTTGGATGCGAGATTTT 1053
DB 1021 ACAGATAAGCTCTTACCTTGGATGCGAGATTTT 1053
```

RESULT 4

```
US-10-172-118-71
; Sequence 71, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
```

```

; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AB036063
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-71

Query Match 99.8%; Score 1051.4; DB 18; Length 4955;
Best Local Similarity 99.9%; Pred. No. 6.8e-294;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 245 ATGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304

QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 305 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364

QY 121 CGGTTTGTCATCTTTCCAAATCAGTACCTTGATATTTGGAAAATGTATAAACAGGCACAG 180
Db 365 CGGTTTGTCATCTTTCCAAATCAGTACCTTGATATTTGGAAAATGTATAAACAGGCACAG 424

QY 181 GCTTCCTTCTGGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAG 240
Db 425 GCTTCCTTCTGGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAG 484

QY 241 CTAAAGCAGATGAGAAGTACTTCACTCTCACTTGTAGCCTTTTTCAGCCAGTGTAT 300
Db 485 CTAAAGCAGATGAGAAGTACTTCACTCTCACTTGTAGCCTTTTTCAGCCAGTGTAT 544

QY 301 GGAATTGTAATGAAAATTTGGTGAGCGCTTTTAGTCAGGAGGTGCAGTTCAGAGGCT 360
Db 545 GGAATTGTAATGAAAATTTGGTGAGCGCTTTTAGTCAGGAGGTGCAGTTCAGAGGCT 604

QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAATGTTTCACTCAGAGATGACAGTTTG 420
Db 605 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAATGTTTCACTCAGAGATGACAGTTTG 664

QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATCAATTTGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATCAATTTGAA 724

QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCTTTCGCGATGATAGCAGATAGAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCTTTCGCGATGATAGCAGATAGAAA 784

QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTTGTCTGTGTAGAAAGAGTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGTGGCTTTTGTCTGTGTAGAAAGAGTTTCTTCTCAGGA 844

QY 601 TCTTTTGTCTATATCTTGGCTTAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
Db 845 TCTTTTGTCTATATCTTGGCTTAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC 904

QY 661 AATGAACATCATCAGCAGAGATGAAGACTTTCATCTGTGACTTTTGTCTGTGCTGTATTTCCAA 720
Db 905 AATGAACATCATCAGCAGAGATGAAGACTTTCATCTGTGACTTTTGTCTGTGCTGTATTTCCAA 964

; APPLICANT: Dai, Hongyue
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-71

Query Match 99.8%; Score 1051.4; DB 19; Length 4955;
Best Local Similarity 99.9%; Pred. No. 6.8e-294;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 245 ATGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304

QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 305 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364

QY 121 CGGTTTGTCATCTTTCCAAATCAGTACCTTGATATTTGGAAAATGTATAAACAGGCACAG 180
Db 365 CGGTTTGTCATCTTTCCAAATCAGTACCTTGATATTTGGAAAATGTATAAACAGGCACAG 424

QY 181 GCTTCCTTCTGGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAG 240
Db 425 GCTTCCTTCTGGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAG 484

QY 241 CTAAAGCAGATGAGAAGTACTTCACTCTCACTTGTAGCCTTTTTCAGCCAGTGTAT 300
Db 485 CTAAAGCAGATGAGAAGTACTTCACTCTCACTTGTAGCCTTTTTCAGCCAGTGTAT 544

QY 301 GGAATTGTAATGAAAATTTGGTGAGCGCTTTTAGTCAGGAGGTGCAGTTCAGAGGCT 360
Db 545 GGAATTGTAATGAAAATTTGGTGAGCGCTTTTAGTCAGGAGGTGCAGTTCAGAGGCT 604

QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAATGTTTCACTCAGAGATGACAGTTTG 420
Db 605 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAATGTTTCACTCAGAGATGACAGTTTG 664

QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATCAATTTGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATCAATTTGAA 724

QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCTTTCGCGATGATAGCAGATAGAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCTTTCGCGATGATAGCAGATAGAAA 784

QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTTGTCTGTGTAGAAAGAGTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGTGGCTTTTGTCTGTGTAGAAAGAGTTTCTTCTCAGGA 844

QY 601 TCTTTTGTCTATATCTTGGCTTAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
Db 845 TCTTTTGTCTATATCTTGGCTTAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC 904

QY 661 AATGAACATCATCAGCAGAGATGAAGACTTTCATCTGTGACTTTTGTCTGTGCTGTATTTCCAA 720
Db 905 AATGAACATCATCAGCAGAGATGAAGACTTTCATCTGTGACTTTTGTCTGTGCTGTATTTCCAA 964
```

Db	425	GCTTCTCTTGGACAGCAGAAAGAGGTGAGCTATTATCAAAAGGATCTCCCTCACTGGAAACAAG	484
Qy	241	CTTAAAGCAGATGAGAAAGTACTTCATCTCTCACATCTTTAGCGCTTTTGTTCAGACCCAGTGAT	300
Db	485	CTTAAAGCAGATGAGAAAGTACTTCATCTCTCACATCTTTAGCGCTTTTGTTCAGACCCAGTGAT	544
Qy	301	GGAAATTCGTAATGAAATTTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT	360
Db	545	GGAAATTCGTAATGAAATTTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT	604
Qy	361	CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTCACAGTTTG	420
Db	605	CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTCACAGTTTG	664
Qy	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTTATTAAATGCAATTGAA	480
Db	665	CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTTATTAAATGCAATTGAA	724
Qy	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCCCTTGCGATCGATAGCAGATAGAAAA	540
Db	725	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCCCTTGCGATCGATAGCAGATAGAAAA	784
Qy	541	TCTACTTTTGGGAAAGAGTGGTGGCCCTTTGCTGCTGTAGAAAGAGATTTTCTTCTCAGGA	600
Db	785	TCTACTTTTGGGAAAGAGTGGTGGCCCTTTGCTGCTGTAGAAAGAGATTTTCTTCTCAGGA	844
Qy	601	TCTTTTGCTGCTATATTCTGGCTAAAGAAAGAGAGGTCTTATGCGCAGACTCACTTTTTTCC	660
Db	845	TCTTTTGCTGCTATATTCTGGCTAAAGAAAGAGAGGTCTTATGCGCAGACTCACTTTTTTCC	904
Qy	661	AATGAACTCATCAGCAGAGATGAAGGACTTCACATGTCACTTTGCTGCTCCCTGATGTTCCAA	720
Db	905	AATGAACTCATCAGCAGAGATGAAGGACTTCACATGTCACTTTGCTGCTCCCTGATGTTCCAA	964
Qy	721	TACTTAGTAAATTAAGCCTTCAGAAAGAAAGGTTCAGGAGATCATTTGTTGATGCTGTCAAA	780
Db	965	TACTTAGTAAATTAAGCCTTCAGAAAGAAAGGTTCAGGAGATCATTTGTTGATGCTGTCAAA	1024
Qy	781	ATTGAGCAGAGTTTTTAAACAGAGCCCTTCCAGCTTGGCCCTCATGGAATGAATTCGATT	840
Db	1025	ATTGAGCAGAGTTTTTAAACAGAGCCCTTCCAGCTTGGCCCTCATGGAATGAATTCGATT	1084
Qy	841	TTGATGAAACAGTACATTGATGTTTGTAGCTGACAGATTTACTTGGAACCTTGGATTTCTCA	900
Db	1085	TTGATGAAACAGTACATTGATGTTTGTAGCTGACAGATTTACTTGGAACCTTGGATTTCTCA	1144
Qy	901	AAGGTTTTTCAGGCAGAAAAATCCCTTTTGATTTTATGGAATAACATTTCTTTAGAAAGAAAA	960
Db	1145	AAGGTTTTTCAGGCAGAAAAATCCCTTTTGATTTTATGGAATAACATTTCTTTAGAAAGAAAA	1204
Qy	961	ACAAATTTCTTTCAGAAACAGATTTTCAGAGTATCAGCGTTTTGTCAGTTATGCGCAAAACC	1020
Db	1205	ACAAATTTCTTTCAGAAACAGATTTTCAGAGTATCAGCGTTTTGTCAGTTATGCGCAAAACC	1264
Qy	1021	ACAGATAAGCTCTTCACTTTGGATGCAATTTT	1053
Db	1265	ACAGATAAGCTCTTCACTTTGGATGCAATTTT	1297

RESULT 6

845	1C	TTTTGCTGCTATATCTCGGCTAAGAGAGAGAGGTC/TATGCGCCAGGACTCATCTTTTCC	904
661	QY	AATGAACCTCATCGACAGAGATGAAGGACTTCTCACTGTGACTTTTGGCTTGCTGTATGTTCCAA	720
905	Db	AATGNACTCATCGACAGAGATGAAGGACTTCACTGTGACTTTGGCTTGCTGTATGTTCCAA	964
721	QY	TACTTAGTAAATAAGCCCTTCAGAGAAAGGGTCAGGAGATCATTTGTTGATCGTGTCAA	780
965	Db	TACTTAGTAAATAAGCCCTTCAGAGAAAGGGTCAGGAGATCATTTGTTGATCGTGTCAA	1024
781	QY	ATTGAGCAGGAGTCTTTTAAACAGAGCCCTTGCCAGTGTGGCCCTCATTTGGAATGAATTCATT	840
1025	Db	ATTGAGCAGGAGTCTTTTAAACAGAGCCCTTGCCAGTGTGGCCCTCATTTGGAATGAATTCATT	1084

841	TTGATGAAACAGTACATTTAGTTGTGTAGCTGACAGATTACTTCTGGAACTTGGATTCTCA	900
1085	TTGATGAAACAGTACATTTAGTTGTGTAGCTGACAGATTACTTCTGGAACTTGGATTCTCA	1144
901	AAGGTTTTTTCAGCGCAGAAAATCCTTTTGTATTTATGGAACAACATTTCTTTAGAACAAAA	960
1145	AAGGTTTTTTCAGCGCAGAAAATCCTTTTGTATTTATGGAACAACATTTCTTTAGAACAAAA	1204
961	ACAAATTTCTTTTCAGAAAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAAAAC	1020
1205	ACAAATTTCTTTTCAGAAAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAAAAC	1264
1021	ACAGATAACGCTCTTCACCTTGGATGCGAGATTTT	1053
1265	ACAGATAACGCTCTTCACCTTGGATGCGAGATTTT	1297

RESULT 7

```

US-10-220-335-514/c
; Sequence 514, Application US/10220335
; Publication No. US20050175607A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-041
; CURRENT APPLICATION NUMBER: US/10/220,335
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 09/664,641
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,807
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/597,707
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/577,409
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 09/515,126
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 688
; SOFTWARE: Custom
; SEQ ID NO 514
; LENGTH: 2482
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-220-335-514

```

Query Match	55.1%	Score 580	DB 24	Length 2482
Best Local Similarity	75.1%	Pred. No. 4.5e-157		
Matches 724	Conservative 0	Mismatches 240	Indels 0	Gaps 0
Qy	89	ATGAAGAGCCACTCCTAAGAAGAGATTCTCGCGGTTTGTCACTTTCCAAATCCAGTACC	148	
Db	2086	AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTCACTTTCCCATCGAGTACC	2027	
Qy	149	CTGATATTTGGAAAATGTATAAACAGGCACAGGCTTCCTTTCTGGACAGCAGAAAGAGTTG	208	
Db	2026	ATGATACTCGCAGATGTATAAGAAGGCAGAGGCTTCCTTTGGACCGCGAGGAGTTG	1967	
Qy	209	ACTTATCAAAAGGATCTCCCTCACTGGAAACAGCTTAAAGCAGATGAGAAAGTACTTCATCT	268	
Db	1966	ACCTCTCCAAGGACATTTCAGCACTGGGAATCCCTTGAAACCCGAGAGCAGATATTTTATAT	1907	
Qy	269	CTCACATCTTAGCCCTTTTTCGACGCCAGTGAATGTAAATGAAAAATTCGTGGAGC	328	
Db	1906	CCCATGTTCTGGCTTCTTTTCGACCAAGCAGATGGCATAGTAAATGAAAAATTCGTGGAGC	1847	
Qy	329	GCCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCTCGCTCTTTCTATGGCTTTCAAATTCCTCA	388	
Db	1846	GAITTAGCCAAGAGTTTCAGATTACAGAGCCCGCTGTTCTATGGCTTCCAAATTCGCCA	1787	
Qy	389	TCGAGAAATGTTCACTCAGAGATGACAGTTTGTCTGATAGACATTTACATCAGAGATCCCA	448	
Db	1786	TGAAAAACATACATTTCTGAAATGTATAGTCTTCTTATTTGACATTTACATAAAAAGATCCCA	1727	

Qy	449	AGAAAGGGAATTTTAAATTAATGCAATTAAGAAACCATAGCCCTATGTTAAAGAAAAAGCAG	508
Db	1726	AAGAAAGGGAATTTTCTCTCAATGCCATTGAAACGATGSCCTGTGTCAAGAAAGAGCAG	1667
Qy	509	ATTGGGCGCTTGGCATGATGATAGAGATAAGAAATCTACTTTTGGGAAAGAGTGTGTGSCCT	568
Db	1666	ACTGGGCGCTTGGCGCTGGATTGGGGAAGAGAGGCTACTATGTGTGAACGTTGTGTAGCCT	1607
Qy	569	TTGCTGCTGTAGAAAGGAGTTTCTCTCTCAGGATCTTTTGTGCTGTATATTCTGGCTAAAGA	628
Db	1606	TTGCTGCTGAGTGAAGGCAATTTCTTTTCCGGTCTTTTTCGGTCTCGATATTCTGGCTCAAGA	1547
Qy	629	AGAGAGTCTTATGCGCAGGACTACTCTTTTTCCAATGAACCTCATCAGCAGAGATGAAGAC	688
Db	1546	AACGAGGACTGATGCTCGGCTCACATTTTCTAATGAACCTTATTAGCAGAGATGAGGGTT	1487
Qy	689	TTCACTGTGACTTTGCTTTGCTGCTGATGTTCCAATACTTCTAGTAAATAAGCCTTCAGAGAAA	748
Db	1486	TACACTGTGATTTTGTCTTGCTGATGTTCAAAACCTGGTACAAACCATCGGAGGAGA	1427
Qy	749	GGGTGCGGAGATCATTTGTTGATGCTGCTCAAAATTTGACGAGGAGTTTTTAAACAGAAACCT	808
Db	1426	GAGTAAGAGAAATAATTATCAATGCTGTTTCGGATAGAACAGGAGTTCTCTCCTGAGGCT	1367
Qy	809	TGCAGTTGGCCTCAITTTGGAATGAATTTGCAATTTTGATGAAACAGTACATTTGAGTTGTAG	868
Db	1366	TGCCTGTGAAGCTCAITTTGGGATGAATTTGCACTCTTAATGAAGCAATACATTTGAGTTTGTG	1307
Qy	869	CTGACAGATTACTTGTGGAACCTGGATTCTCAAGAGTTTTTCAGGCGAGAAATCCTTTTG	928
Db	1306	CAGACAGACTTATGCTGGAACTGGGTTTATAGCAAGGTTTTCAGAGTAGAGAACCCATTG	1247
Qy	929	ATTTTATGGAAGCAATTTCTTTAGAGGAAAAACAAATTTCTTTTGAAAAACGAGTTTCAG	988
Db	1246	ACTTTATGAGAGATATTTTCACTGGAAGAAAGACTAATCTTTTGAAGAGAGTAGCG	1187
Qy	989	AGTATACGCGTTTTCAGTTATGCGAGAAACCAACAGATAACGTTCTTCCCTTGATGAG	1048
Db	1186	AGTATCAGAGGATGGGAGTGATGTCAAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTG	1127
Qy	1049	ATTT 1052	
Db	1126	ACTT 1123	
RESULT 8			
US-09-954-456-724			
; Sequence 724, Application US/09954456			
; Patent No. US20020115057A1			
; GENERAL INFORMATION:			
; APPLICANT: Young, Paul			
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents U			
; TITLE OF INVENTION: Sets			
; FILE REFERENCE: 689290-76			
; CURRENT APPLICATION NUMBER: US/09/954,456			
; CURRENT FILING DATE: 2001-09-18			
; PRIOR APPLICATION NUMBER: US/60/233,617			
; PRIOR FILING DATE: 2000-09-18			
; PRIOR APPLICATION NUMBER: US/60/234,052			
; PRIOR FILING DATE: 2000-09-20			
; PRIOR APPLICATION NUMBER: US/60/234,923			
; PRIOR FILING DATE: 2000-09-25			
; PRIOR APPLICATION NUMBER: US/60/235,134			
; PRIOR FILING DATE: 2000-09-25			
; PRIOR APPLICATION NUMBER: US/60/235,637			
; PRIOR FILING DATE: 2000-09-26			
; PRIOR APPLICATION NUMBER: US/60/235,638			
; PRIOR FILING DATE: 2000-09-26			
; PRIOR APPLICATION NUMBER: US/60/235,711			
; PRIOR FILING DATE: 2000-09-27			
; PRIOR APPLICATION NUMBER: US/60/235,720			
; PRIOR FILING DATE: 2000-09-27			
; PRIOR APPLICATION NUMBER: US/60/235,840			

; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,863
 ; PRIOR FILING DATE: 2000-09-27
 ; NUMBER OF SEQ ID NOS: 2276
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 724
 ; LENGTH: 2500
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-954-456-724

Query Match 55.1%; Score 580; DB 9; Length 2500;
 Best Local Similarity 75.1%; Pred. No. 4.6e-157;
 Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY	89	ATGAGAGCCACTCTCTAAGAAAGAGTTCTCGCGGGTTTGTCTATCTTTCCAAATCCAGTACC	148
DB	397	AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTCTATCTTTCCCATCGAGTACC	456
QY	149	CTGATATTTGGAAATGTATACAGGCACAGCTTCTCTGACAGCAGAGAGGTTG	208
DB	457	ATGATATCTGGCAGATGATAGAGCGAGAGCTTCTTTGGACCGCGAGAGGTTG	516
QY	209	ACTTATCAAGAGATCTCCCTCACTGGAAACAGCTTAAAGCAGATGAGAGTACTTCTATCT	268
DB	517	ACCTCTCCAAGACATTCAGCACTGGGAATCCCTGAAACCCGAGAGAGATATTTTATAT	576
QY	269	CTCAGATCTTAGCCTTTTTCAGCCAGTGTATGGAATGTAAATGAAAAATTTGGTGGAGC	328
DB	577	CCCATGTTCTGGCTTTCTTTGACAAAGCGATGGCATAGTAAATGAAAACTTTGGTGGAGC	636
QY	329	GCTTTAGTCAGAGGTGCGAGGTTCCAGAGCTCGCTGTTCTATGCTTTTCAAAATCTCA	388
DB	637	GATTTAGCCAAAGTTTCAATATGATGAGGAGGCTGTTCTATGCTTTTCAAAATTTGCCA	696
QY	389	TCGAGATCTTCACTCAGAGATGTACAGTTTCTGATGACACTTACATCAGAGATCCCA	448
DB	697	TGGAACACATATCTGAAATGTATAGTCTTTTATGACACTTACATAAAGATCCCA	756
QY	449	AGAAAGGGAAATTTTATTAATGCAATGAAACCATGCCCTATGTTAAGAAAAAGCAG	508
DB	757	AAAGAAAGGAAATTTCTTCAATGCCATTGAAACCATGCCCTGTTGTCAGAAAGAGGCAG	816
QY	509	ATTGGCCCTTGGATGATGATGAGATAGAAATCTACTTTTGGGGAAAGAGTGGTGCCT	568
DB	817	ACTGGCCCTTGGCTGGATGAGGAGCAAGAGGCTACCTATGTTGAAACGCTGTTGAGCCT	876
QY	569	TTGCTGCTGTAGAAGGATTTTCTTCTCAGGATCTTTTCTGCTGATATTTCTGCTAAACA	628
DB	877	TTGCTGCTGTAGAAGGATTTTCTTCTTCTGCTGATATTTCTGCTGATATTTCTGCTCAAGA	936
QY	629	AGAGAGGCTTATGTCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAGGAC	688
DB	937	RACGAGGACTGATGCTGCGCTCACATTTTCTAATGAACTTATAGCAGAGATGAGGTT	996
QY	689	TTCACTGTGACTTTGCTGCTGATGTTCCAAATCTACTTAAATAGACCTTCAGAGAAA	748
DB	997	TACACTGTGATTTGCTGCTGATGTTCAAAACACCTGTGATACACAAACCATCGGAGGAG	1056
QY	749	GGGTCAGGAGATCATTTGATGCTGTCAAAATTTGAGCAGAGGTTTAAACAGAGCCT	808
DB	1057	GAGTAAAGAAATTAATCAATGCTGTTGCGATGAAACAGGAGTTCCTCACTGAGGCT	1116
QY	809	TGCCAGTTGGCTCATTTGGAATGAATTCATTTTGAATGAAACAGTACATTTGATTTGTAG	868
DB	1117	TGCTGTGAAGCTCATTTGGATGAATTCGACTCTAATGAGCAATACATTTGATTTGTG	1176
QY	869	CTGACAGATTTACTTGTGAACTTGTGATTTCTCAAAAGTTTTCAGGCAAAAAATTCCTTTG	928
DB	1177	CAGACAGACTTATGCTGGAATCTGGGTTTATAGCAAGGTTTTCAGAGTATGAGAACCCATTG	1236
QY	929	ATTTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTTGAACACAGTTTCAG	988

DB 1237 ACTTTATGGAGAATATTTTCACTGGAAGAAAGACTAACTTCTTTGAGAAGAGTAGGCG 1296
 QY 989 AGTATCAGCGTTTTCAGTTATGCGACAAAACCAAGATAAAGTCTTTACCTTGGATGAC 1048
 DB 1297 AGTATCAGAGGATGGGAGTGTATGTAAGTCCACAGAGAAATTTCTTTTACCTTGGATGCTG 1356
 QY 1049 ATTT 1052
 DB 1357 ACTT 1360

RESULT 9

US-09-954-456-1169
 ; Sequence 1169, Application US/09954456
 ; Patent No. US20020115057A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Young, Paul
 ; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can
 ; TITLE OF INVENTION: Sets
 ; FILE REFERENCE: 689290-76
 ; CURRENT APPLICATION NUMBER: US/09/954,456
 ; CURRENT FILING DATE: 2001-09-18
 ; PRIOR APPLICATION NUMBER: US/60/233,617
 ; PRIOR FILING DATE: 2000-09-18
 ; PRIOR APPLICATION NUMBER: US/60/234,052
 ; PRIOR FILING DATE: 2000-09-20
 ; PRIOR APPLICATION NUMBER: US/60/234,923
 ; PRIOR FILING DATE: 2000-09-25
 ; PRIOR APPLICATION NUMBER: US/60/235,134
 ; PRIOR FILING DATE: 2000-09-25
 ; PRIOR APPLICATION NUMBER: US/60/235,637
 ; PRIOR FILING DATE: 2000-09-26
 ; PRIOR APPLICATION NUMBER: US/60/235,638
 ; PRIOR FILING DATE: 2000-09-26
 ; PRIOR APPLICATION NUMBER: US/60/235,711
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,720
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,840
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,863
 ; PRIOR FILING DATE: 2000-09-27
 ; NUMBER OF SEQ ID NOS: 2276
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 1169
 ; LENGTH: 2500
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-954-456-1169

Query Match 55.1%; Score 580; DB 9; Length 2500;
 Best Local Similarity 75.1%; Pred. No. 4.6e-157;
 Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY	89	ATGAGAGCCACTCTCTAAGAAAGAGTTCTCGCGGGTTTGTCTATCTTTCCAAATCCAGTACC	148
DB	397	AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTCTATCTTTCCCATCGAGTACC	456
QY	149	CTGATATTTGGAAATGTATACAGGCACAGCTTCTCTGACAGCAGAGAGGTTG	208
DB	457	ATGATATCTGGCAGATGATAGAGCGAGAGCTTCTTTGGACCGCGAGAGGTTG	516
QY	209	ACTTATCAAGAGATCTCCCTCACTGGAAACAGCTTAAAGCAGATGAGAGTACTTCTATCT	268
DB	517	ACCTCTCCAAGACATTCAGCACTGGGAATCCCTGAAACCCGAGAGAGATATTTTATAT	576
QY	269	CTCAGATCTTAGCCTTTTTCAGCCAGTGTATGGAATGTAAATGAAAAATTTGGTGGAGC	328
DB	577	CCCATGTTCTGGCTTTCTTTGACAAAGCGATGGCATAGTAAATGAAAACTTTGGTGGAGC	636
QY	329	GCTTTAGTCAGAGGTGCGAGGTTCCAGAGCTCGCTGTTCTATGCTTTTCAAAATCTCA	388
DB	637	GATTTAGCCAAAGTTTCAATATGATGAGGAGGCTGTTCTATGCTTTTCAAAATTTGCCA	696

```
QY 389 TCGAGAAATGTTCACTCAGAGATGTACAGTTCTGCTCATAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAACATACATCTGAAATGATAGTCTTCTTATGACACTTACATTAAGATCCCA 756
QY 449 AGAAAGGGAAATTTTATTAATGCAATGGAACCATGCCCTATGTTAAGRAAAAGCAG 508
Db 757 AAGAAAGGGAAATTTCTCTCAATGCCAATGAAACGATGCTGTGTCAAGRAAGAGCAG 816
QY 509 ATTGGGCTTGGGATGGATAGACATAGAAATCTACTTTTGGGGAAGAGATGGTGGCT 568
Db 817 ACTGGGCTTGGGATGGGATGGGACAAAGAGGCTACCTATGGTGAACGCTGTGTAGCCT 876
QY 569 TTGCTGCTGTAGAGAGATTTCTCTCAGAGATCTTTGCTGCTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGCTGAGAGAGATTTCTCTTCCGGTCTTTTGGTGCATATTTCTGGCTCAAGA 936
QY 629 AGAGAGGCTTTATGCGAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCTGGCTCACATTTTCTAATGAACTTATTAGCAGAGATGAGGTT 996
QY 689 TTCACTGTGACTTTGCTTGGCTGATGTTCCAAATGCTTCTGCTATATTTCTGGCTAAAGA 748
Db 997 TACACTGTGATTTTGTCTTGGCTGATGTTTCAACACCTGGTACACAAACCATCGAGGAGA 1056
QY 749 GGGTCAGGAGATCAATTTGTTGATGCTGTCAAAATTTGAGCAGGAGATTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATAATTATCAATGCTGTTCGGATAGAACAGGAGTTCTCTACTGAGCCT 1116
QY 809 TGCCAGTTGGGCTCATTTGGAATGAAATGCAATTTGATGAAACAGTACATTTGATGTTGTAG 868
Db 1117 TGCCGTGGAAGCTCATTTGGGATGAAATGCACTCTAATGAAGCAATACATTTGATTTGTG 1176
QY 869 CTGACAGATTAATTTGTTGGAATTTGGAATTTCAAGGTTTTCAGGCAAGAAATTCCTTTG 928
Db 1177 CAGACAGACTTATGCTGGAATGCTGTTTTCAGGATGAAACAGAGTTCTCTACTGAGGCT 1236
QY 929 ATTTTATGGAACAAATTTCTTTAGAGGAACAAATTTCTTTGAGAAACAGTTTTCAG 988
Db 1237 ACTTTATGGAATAATTTTCTGGAAGGAAGAACTAATCTTTTGAAGAGAGATGAGCG 1296
QY 989 AGTATCAGCTTTTGCAGTTATGCGAGAAACACAGATAAAGCTTTCACTTTGATGCGAG 1048
Db 1297 AGTATCAGAGGATGGAGTGTCAAGTCCAAACAGAGAAATTTCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 10
US-09-954-456-1827
; Sequence 1827, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; FILE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
```

```
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 1827
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1827

Query Match 55.1%; Score 580; DB 9; Length 2500;
Best Local Similarity 75.1%; Pred. No. 4.6e-157;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCACTCTTAAGAAAGAGTTCTGCGCGGTTTGTTCATCTTTCCCATCCAGTACC 148
Db 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTTCATCTTTCCCATCCAGTACC 456
QY 149 CTGATATTTGGAAATGTATAACAGGCACAGGCTTCTCTTGACACAGAGAGAGGTTG 208
Db 457 ATGATATCTGGCAGATGTATAGAAGGCAGAGGCTTCTTTTGGACCCCGAGGAGGTTG 516
QY 209 ACTTATCAAAAGGATCTCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTCATCT 268
Db 517 ACCTCTCCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
QY 269 CTCACATCTTAGGCTTTTTCGAGCCAGTGTGGAATTTGTAATGAAAAATTTGGTGGAGC 328
Db 577 CCATGTTCTGCGCTTTCTTTGACGCAAGCGATAGTAAATGAAAACTTGGTGGAGC 636
QY 329 GCTTTAGTCAGAGGTCAGGTTTCAGAGGCTCGCTGTTTCTATGGCTTTCAAAATCTCA 388
Db 637 GATTTAGCCAGAGTTTCAGATTTACAGAGCCGCTGTTTCTATGGCTTTCAAAATCTCCA 696
QY 389 TCGAGAAATTTCACTCAGAGATGTACAGTTTGTGTATAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAAAACATACATTTCTGAAAATGATAGTCTTCTTATTTGACACTTACATAAAGATCCCA 756
QY 449 AGAAAAAGGAATTTTATTTAATGCAATTTGAAACCATGCCCTATGTTAAGAAAAAGCAG 508
Db 757 AAGAAAGGGAAATTTCTCTCAATGCCATTTGAAACGATGCGCTTGTGTCAAGAAAGAGCAG 816
QY 509 ATTGGGCTTGGATGCGATAGCAGATAGAAAAATCTACTTTTGGGGAAGAGTGGTGGCT 568
Db 817 ACTGGGCTTGGCTGGATTTGGGACAAAGAGGCTACCTAATGTTGAACGTTGTTGTAGCCT 876
QY 569 TTGCTGCTGTAGAAGGAGTTTCTCTCTCAGGATCTTTTGTGCTGTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGCTGTAGAAGGCAATTTCTTTCCGGTCTTTTGTGCTGATATTTCTGGCTCAAGA 936
QY 629 AGAGAGGCTTTATGCGAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCTGGCTCACATTTTCTAATGAACTTATTAGCAGAGATGAGGTT 996
QY 689 TTCACTGTGACTTTGCTTGGCTGATGTTTCAATGCTTAGTAAATAGCCTTTCAAGAGAAA 748
Db 997 TACACTGTGATTTTGTCTTGGCTGATGTTTCAACACCTGGTACACAAACCATCGAGGAGA 1056
QY 749 GGGTCAGGAGATCAATTTGTTGATGCTGTCAAAATTTGAGCAGGAGATTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATAATTATCAATGCTGTTCGGATAGAACAGGAGTTCTCTACTGAGCCT 1116
QY 809 TGCCAGTTGGCCTCATTTGGAATGAAATGCAATTTGATGAAACAGTACATTTGATGTTGTAG 868
Db 1117 TGCCGTGGAAGCTCATTTGGGATGAAATGCACTCTAATGAAGCAATACATTTGATTTGTG 1176
QY 869 CTGACAGATTAATTTGTTGGAATTTGGAATTTCAAGGTTTTCAGGCAAGAAATTCCTTTG 928
Db 1177 CAGACAGACTTATGCTGGAATGCTGTTTTCAGGATGAAACAGAGTTCTCTACTGAGGCT 1236
QY 929 ATTTTATGGAACAAATTTCTTTAGAGGAACAAATTTCTTTGAGAAACAGTTTTCAG 988
Db 1237 ACTTTATGGAATAATTTTCTGGAAGGAAGAACTAATCTTTTGAAGAGAGATGAGCG 1296
QY 989 AGTATCAGCTTTTGCAGTTATGCGAGAAACACAGATAAAGCTTTCACTTTGATGCGAG 1048
Db 1297 AGTATCAGAGGATGGAGTGTCAAGTCCAAACAGAGAAATTTCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 10
US-09-954-456-1827
; Sequence 1827, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; FILE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
```

Db 1177 CAGACAGACTTATGCTGGAACTGGGTTTATAGCAAGGTTTTCAGAGTAGAAGACCCATTG 1236
Qy 929 ATTTATGAAAAACATTTCTTTAGAAAGGAAAAAATAATTTCTTTTGAGAAACAGAGTTTCAG 988
Db 1237 ACTTTATGAGAAATATTTCACTGGAAGGAAGAGTAACTTTCTTTTGAGAAAGAGTAGGGC 1296
Qy 989 AGTATCAGCGGTTTTCAGTTATGGCAGAAACCAACAGATAACGCTTTCACCTTCGATGCCAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGATGCAAGTCCAAACAGAGAAATTTTATACCTTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 11

US-10-641-643-1370
; Sequence 1370, Application US/10641643
; Publication No. US20040077003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; Susan G. Stuart
; Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
; GENE EXPRESSION

NUMBER OF SEQUENCES: 1508
CORRESPONDENCE ADDRESS:
ADDRESS: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/641,643

FILING DATE: 14-Aug-2003

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: <Unknown>

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0001 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 1370:

SEQUENCE CHARACTERISTICS:

LENGTH: 2500 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GENBANK

CLONE: g36154

SEQUENCE DESCRIPTION: SEQ ID NO: 1370 :

US-10-641-643-1370

Query Match 55.1%; Score 580; DB 19; Length 2500;

Best Local Similarity 75.1%; Pred. No. 4.6e-157;

Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

Qy 89 ATGAAGACCACTCTCTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTTTCAATCCAGTACC 148

Db 397 AGGATGACGCGCTGCTGAGAGAAACCCCGCGCTTTGTCTATCTTCCCATCGAGTACC 456

Qy 149 CTGATATTTGGAAAAATGTATAAACAGGCACAGAGCTTCTCTTGGACAGCAGAGAGGTTG 208
Db 457 ATGATATCTGGCAGATGTATAAGAGGAGCAGAGGCTTCTCTTTGGACCCCGAGGAGGTTG 516
Qy 209 ACTTATCAAAAGGATCTCCCTCAGCTGGAACAAAGCTTTAAAGCAGATGAGAGTACTTCATCT 268
Db 517 ACCTCTCAAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
Qy 269 CTCACATCTTAGCCTTTTTCAGCCAGTGTAAATGTAATCAAAATTTGGTGGAGC 328
Db 577 CCCATGTTGGGCTTCTTTTCAGCAAGCGATGTCATAGTAAATGAAACATTTGGTGGAGC 636
Qy 329 GCTTTAGTCAGGAGTGCGAGTTCAGAGGCTCGCTGTTTCTATGGGCTTCAAAATTTCTCA 388
Db 637 GATTTAGCCAAAGATTCAGATTACAGAGCCGCTGTTTCTATGGCTTCCAAATTTGCCA 696
Qy 389 TCGAGAAATGTTCACTCAGAGATGTACAGTTTGTGTGATAGACACTTATACATCAGAGATCCCA 448
Db 697 TGGAAAAACATACATTTCTGAAATGTATAGTCTTCTTATTGACACTTACATAAAAGATCCCA 756
Qy 449 AGAAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCCCTATGTTAAGAAAAAGAGC 508
Db 757 AAGAAAAGGGAATTTCTCTTCAATGCCATTTGAAACGATGCCCTGTGTCAAGAAAGAGGAG 816
Qy 509 ATTGGGCTTGGCAGTAGCAGATAGAAAACTCTCTTTTGGGGAAGAGTGGTGGCT 568
Db 817 ACTGGGCTTGGCCTGGATTGGGACAAAGGCTACTATGGTGAACGTGTGTGTAGGCT 876
Qy 569 TTGCTGCTGTAAGAGGAGTTTCTTCTCAGGATCTTTTGTGCTGTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGCAAGTGAAGGCAATTTCTTTTCCGCTTCTTTTGGCTCGATATTTCTGGCTCAAGA 936
Qy 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCAAATGAACCTCATCAGCAGAGATGAAGAC 688
Db 937 AACGAGGACTGATGCCCTGGCTCACATTTTCTAATGAACCTTATAGCAGAGATGAGGTT 996
Qy 689 TTCACGTGACTTTGCTTGGCTGCTGATGTTCCAATACTTTAGTAAATAAGCCTTCAGAGAA 748
Db 997 TACACTGTGATTTTGTCTGCTGATGTTCAAAACACCTGGTACACAAACCATCGGAGGAGA 1056
Qy 749 GGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTTGAGCAGGAGTCTTTTAAACAGAGCT 808
Db 1057 GAGTAAGAGAAATTAATTAATCAATGCTTTCGGATAGAACAGGAGTTCTCTACTGAGGCT 1116
Qy 809 TGCCAGTTGCCCTCATTTGGAATGAATGCAATTTGATGAAACAGTACATTCAGTTGTAG 868
Db 1117 TGCTGTGAAGCTCAITGGGATGAATGCACTCTAATGAAGCAATACATTCAGTTGTGG 1176
Qy 869 CTGACAGATTACTTTGGGACTTGGATTCTCAAGGTTTTTCAGGCGAGAAATCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACTGGGTTTTAGCAAGGTTTTTTCAGAGTAGAGAACCCATTG 1236
Qy 929 ATTTTATGAAAAACATTTCTTTTAAAGGAAAAACAAATTTCTTTTGAAAAACAGTTTCAG 988
Db 1237 ACTTTATGAGATATTTTCTCGAAGGAAGAGTAACTTCTTTTGAGAGAGAGTAGCGC 1296
Qy 989 AGTATCAGCGTTTTCAGTTATGGCAGAAACCAAGATAACGTTCTTCCCTTGGATGAGC 1048
Db 1297 AGTATCAGAGGATGGGAGTGTCAAGTCCAAACAGAGAAATTTCTTTTACCTTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 12

US-10-733-878-458
; Sequence 458, Application US/10733878
; Publication No. US20040224408A1
; GENERAL INFORMATION:
; APPLICANT: Jean-Philippe Girard
; APPLICANT: Francois Amalric
; APPLICANT: Myriam Rousseigne

```

; APPLICANT: Thomas Clouaire
; TITLE OF INVENTION: THAP PROTEINS AS NUCLEAR RECEPTORS FOR
; TITLE OF INVENTION: CHEMOKINES AND ROLES IN TRANSCRIPTIONAL REGULATION, CELL
; TITLE OF INVENTION: PROLIFERATION AND CELL DIFFERENTIATION
; FILE REFERENCE: BIOBANK.012A
; CURRENT APPLICATION NUMBER: US/10/733.878
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 60/432699
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/485027
; PRIOR FILING DATE: 2003-07-03
; NUMBER OF SEQ ID NOS: 535
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 458
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-733-878-458

```

```

Query Match 55.1%; Score 580; DB 21; Length 2500;
Best Local Similarity 75.1%; Pred. No. 4.6e-157; Indels 0; Gaps 0;
Matches 724; Conservative 0; Mismatches 240;

QY 89 ATGAAGGCCACTCTCTAAGAAAGAGATTCTCGCGGTTTGTGTCATCTTTCCAAATCCAGTACC 148
DB 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTGTCATCTTCCCATCGAGTACC 456

QY 149 CTGATATTTGAAATATGATATAACAGGCACAGGCTTCTTGGACAGAGAGAGGTTG 208
DB 457 ATGATATCTGCGAGATGATAGAGGCGAGGCTTCTTTTGGACCGCGGAGGTTG 516

QY 209 ACTTATCAAGGATCTCCCTCAGTGGAAACAGCTTAAAGCAGATGAGAGTACTTCATCT 268
DB 517 ACCTTCAAGGACATTCAGCAGTGGGAATCCCTGAAACCGGAGAGAGATATTTTATAT 576

QY 269 CTCACATCTTAGCCCTTTTTCAGCGCAGTGTGAATTTGAAATTTGTTGGAGC 328
DB 577 CCCATGTTCTGCTTCTTTCAGCAAGGATGGATAGTAATGAAATTTGTTGGAGC 636

QY 329 GCTTTAGTACAGAGGTGAGGTTTCAGAGGCTGCTGTTTCTATGGCTTTCAAAATCTCA 388
DB 637 GATTAGCAAGAGTTTCAAGATTAAGAAAGCGCGCTGTTTCTATGGCTTCAAAATTCGA 696

QY 389 TCGAGATGTTCACTCAGAGATGTACAGTTTGTCTATAGACACTTACATCAGAGATCCCA 448
DB 697 TGGAAACATACATCTCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 756

QY 449 AGAAAGGGAATTTTATTAATGCAATGAAACCATGCGCTTATGTTAAGAAAGGAGCAG 508
DB 757 AAGAAAGGGAATTTCTTCTCAATGCAATGAAACCATGCGCTTGTCTAAGAAAGGAGCAG 816

QY 509 ATTGGGCTTCGATGAGATAGAGATAGAAATCTATTTTGGGAAAGAGTGGGCT 568
DB 817 ACTGGGCTTCGCTGGATTTGGGACAAAGAGGCTACCTATGTTGAACGTTGTTAGCCT 876

QY 569 TTGCTGCTGTAGAGGAGTTTCTTCTCAGAGTCTTTTGTCTATATCTTGGCTAAAGA 628
DB 877 TTGCTGCTGTAGAGGAGTTTCTTCTTCCGCTTCTTTTCCGCTGATATCTTGGCTCAAGA 936

QY 629 AGAGAGGCTTATGCGAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGAC 688
DB 937 AACGAGGACTGATGCTGCGCTTCACATTTTCTAATGAACTTATAGCAGAGATGGGTT 996

QY 689 TTCACTGTGACTTTGCTTTCGCTGATGTTTCCAAATCTTATGTAATTAAGGCTTTCAGAGAAA 748
DB 997 TACACTGTGATTTTGTCTTGGCTGATGTTTCAACACCTGTACACAAACCATCGGAGGAGA 1056

QY 749 GGGTCAGGAGATCATTTGTTGATGCTGCAAAATTTGAGCAGAGGTTTAAACAGAGGCT 808
DB 1057 GAGTAAGAGAAATAATTAATCAATGCTGTTTCGATAGAACAGGAGTTTCTCTCACTGAGGCT 1116

QY 809 TGCCAGTTGGCCTCATTTGGAATGAAATTTGATGAAACAGATGAGATGAGTTGTAG 868

```

```

DB 1117 TGCCTGTGAAGCTCATTTGGATGAATTGCACCTCTATGAAGCAATACATTGAGTTTGTGG 1176
QY 869 CTGACAGATTACTTGTGGAACTTGAATCTCAAGGTTTTCAGGCGAGAAAATCCTTTTG 928
DB 1177 CAGACAGACTTATGCTGGAACCTGGGTTTTCAGAAAGGTTTTCAGAGTAGAAGACCCATTG 1236
QY 929 ATTTTATGAAAAACATTTCTTTAGAAAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
DB 1237 ACTTTATGGAGATATTTTCACTGGAAGGAAAGACTTAACCTTTTGAAGAGAGTAGGCG 1296
QY 989 AGTATCAGCGTTTTCAGATTATGGCAGAAACACAGATTAACGCTTTCACCTTGGATGACAG 1048
DB 1297 AGTATCAGAGGATGGAGTGTCAAGTCCAACAGAGAAATCTTTTACCTTGGATGCTG 1356
QY 1049 ATTTT 1052
DB 1357 ACTT 1360

```

```

RESULT 13
US-10-843-641A-3751
; Sequence 3751, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; TITLE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 847
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3751
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-3751

```

```

Query Match 55.1%; Score 580; DB 22; Length 2500;
Best Local Similarity 75.1%; Pred. No. 4.6e-157; Indels 0; Gaps 0;
Matches 724; Conservative 0; Mismatches 240;

QY 89 ATGAAGGCCACTCTCTAAGAAAGAGATTCTCGCGGTTTGTGTCATCTTTCCAAATCCAGTACC 148
DB 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTGTCATCTTCCCATCGAGTACC 456

QY 149 CTGATATTTGAAATATGATATAACAGGCACAGGCTTCTTTCGACAGAGAGAGGTTG 208
DB 457 ATGATATCTGCGAGATGATAGAGGCGAGGCTTCTTTTGGACCGCGGAGGAGGTTG 516

QY 209 ACTTATCAAGGATCTCCCTCAGTGGAAACAGGTTTAAAGCAGATGAGAGTACTTCATCT 268

```

Db 517 ACCTCTCCAAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
QY 269 CTCACATCTTAGCCTTTTTCAGCCAGTGATGGAATCTTAATGAAATTTGGTGGAGC 328
Db 577 CCATGTTCTGCTTCTTTGAGCAAGGATGGATAGTAATGAAATCTTTGGTGGAGC 636
QY 329 GCTTTAGTCAGGAGGTGCAAGGTTTCAGAGGCTCGCTGTTTCTATGCTTTTCAAAATCTCA 388
Db 637 GATTTAGCCCAAGATTCAGATTACAGAAGCCGCTGTTCTATGGCTTCCAAATTTGCCA 696
QY 389 TCGAAGATGTTCACTCAGAGATGTACAGTTGCTGATACACTTACATCAGAGATCCCA 448
Db 697 TCGAAGAACATACATTCGAAATGTAAGTCTTCTTATTGACACTTACATAAAAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTATTAATGCAATGGAACCATGCCCTATGTTAAGAAAAGAGC 508
Db 757 AAGAAGGGAATTTCTCTCAATGCAATGGAACCATGCCCTGTTGTTCAAGAAGAGGCGAG 816
QY 509 ATTGGGCTTGGATGGATAGCAGATAGAAAATCTACTTTTGGGAAAGAGTGGTGGCCT 568
Db 817 ACTGGGCTTGGCTGGATGGGACAAAGAGCTACCTATGTTGACGTTGTTAGCCT 876
QY 569 TTGCTGCTGTAAGAGGTTTCTCTCAGGATCTTTTTCAGGATCTTTTCTGCTATATTCTGCTAAAGA 628
Db 877 TTGCTGCACTGGAAGGCAATTTCTTTCCGGTCTTTTTCGGTCGATATTTCTGCTCAAGA 936
QY 629 AGAGAGTCTTATGTCAGGACTCACTTTTCCAAATGAACTATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCTGCGCTTCACATTTTCTTAATGAATTTATTAGCAGAGATGAGGGTT 996
QY 689 TTCCTGTGACTTTCCTGCTGATGTTTCCAAATGAACTATGTTAAGAAAAGAGC 748
Db 997 TACACTGTGATTTTGGCTGCTGATGTTTCCAAATGAACTATGTTAAGAAAAGAGC 1056
QY 749 GGGTCAGGAGATCAATGTTGATGCTGCTCAAAATTTAGCAGGAGTGTTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAATAATTAATCAATGCTGTTTCGGATAGACAGAGTTCCTACTGAGCCT 1116
QY 809 TGCCAGTTGGCTCATTTGGAATGAATTTGCAATTTGATGTAAGAACAGTACATTTGATGTTGATG 868
Db 1117 TGCTGCTGAAAGCTCATTTGGATGAATTTGCACTTAATGAAGCAATACATTTGATTTGTTGG 1176
QY 869 CTGACAGATTAATGCTGAACTTTGCAATTTGCAAGGTTTTCAGGAGGAAATCTTTTG 928
Db 1177 CAGACAGACTTAATGCTGAACTTTGCAAGGTTTTCAGAGGTAGAGAACCCATTTTG 1236
QY 929 ATTTTATGGAACAAATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
Db 1237 ACTTTATGGAGATATTTCACTGGAAGGAAAGACTTAATCTTTTGAAGAGAGTAGGCG 1296
QY 989 AGTATCAGGTTTTCAGTTATGCGAGAAACCAAGATTAAGCTTTCACTTTGGATGCGAG 1048
Db 1297 AGTATCAGAGGATGGAGTGATGTTCAAGTCCCAACAGAGAAATTTCTTTTACCTTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 14
US-10-843-641A-4196
; Sequence 4196, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531

; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4196
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-4196

Query Match 55.1%; Score 580; DB 22; Length 2500;
Best Local Similarity 75.1%; Pred. No. 4.6e-157;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 89 ATGAAGGACCTCTTAAGAAGAGTTCCTCGCGGTTTGTCTATCTTTCCATCCAGTACC 148
Db 397 AGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTCTATCTTTCCCATCCAGTACC 456
QY 149 CTGATATTTGGAATAATGTATAAAGCAGCAGAGTTCCTCTTGACAGCAGAGAGGTTG 208
Db 457 ATGATATCTGGCAGATGTATAAAGGCGCAGAGGCTTCTTTTGACCGCGAGAGGTTG 516
QY 209 ACTTATCAAGGATCTCCCTCACTGGAAACAGCTTAAAGCAGATGAGAAGTACTTCATCT 268
Db 517 ACCTCTCCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
QY 269 CTCACATCTTAGCCTTTTTCAGCCAGTGTGAATTTAAATGAAAAATTTGGTGGAGC 328
Db 577 CCATGCTTCTGGCTTTCTTTTGACAGCAGCGATAGTAAATGAAAACTTTGGTGGAGC 636
QY 329 GCTTTAGTCAGGAGGTGCGAGTTTCAGAGGCTCGCTCTTTCTATGCGCTTTCAAAATCTCA 388
Db 637 GATTTAGCCCAAGATTCAGATTACAGAAGCCGCTGTTCTATGGCTTCCAAATTTGCCA 696
QY 389 TCGAAGATGTTCACTCAGAGATGTACAGTTTGTGTAGTACACTTACATCAGAGATCCCA 448
Db 697 TCGAAGAACATACATTCGAAATGTAAGTCTTCTTATTGACACTTACATAAAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTATTAATGCAATTTGAACCATGCCCTATGTTAAGAAAAGAGC 508
Db 757 AAGAAGGGAATTTCTCTCAATGCCATTTGAAACGATGCCCTGTTGTCAAGAAAGAGGCGAG 816
QY 509 ATTGGGCTTGGATGGATAGCAGATAGAAAATCTACTTTTGGGAAAGAGTGGTGGCCT 568
Db 817 ACTGGGCTTGGCTGGATTTGGGACAAAGAGGCTTACCTATGTTGACGTTGTTGAGCCT 876
QY 569 TTGCTGCTGTAAGGAGTTCCTCTCAGGATCTTTTGTGCTGCTATATTCTGGCTAAAGA 628
Db 877 TTGCTGCTGTAAGGAGCAATTTCTTTTCCGGTCTTTTTCGGTCGATTTCTGGCTCAAGA 936
QY 629 AGAGAGTCTTATGCGAGGACTCACTTTTCCAAATGAACTATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCTGCGCTTCACATTTTCTAATGAACCTTATTAGCAGAGATGAGGGTT 996
QY 689 TTCCTGTGACTTTCCTGCTGATGTTTCCAAATGAACTATGTTAAGAAAAGAGC 748
Db 997 TACACTGTGATTTTGGCTGCTGATGTTTCCAAACACCTGGTACACAAACCATTCGGAGGAGA 1056

QY	749	GGGTCAGGAGATCATTTGTCGTCAAAATTTGACGAGGTTTTTAACAGAGCCT	808
Db	1057	GAGTAAGAGAAATTAATTAATGCTGTCGATAGAAACAGAGTTCCTCACTGAGGCT	1116
QY	809	TGCCAGTTGGCCTCATTTGGAATGCAATTTTGATGAAACAGTACATTTGATG	868
Db	1117	TGCCTGTGAAGCTCATTTGGGATGATTCACCTCTAATGAAGCAATACATTTG	1176
QY	869	CTGACAGATTAATTTGTTGGAATTTGCAATTTTCAAGGTTTTTTCAGGAGAAATC	928
Db	1177	CAGACAGACTTATGCTGGAACCTGGGTTTTTAGCAAGTTTTTTCAGAGTAGAG	1236
QY	929	ATTTTATGGAACATTTCTTTAGAGCAAAACAAATTTCTTTAGAAACGAGTTTC	988
Db	1237	ACTTTATGGAATATTTTCTGGAAGAAAGAAAGTAACTTTCTTTTGAAGAGAT	1296
QY	989	AGTATCAGCGTTTTCAGTTATGGCAGAAACCAAGATAACCGTCTTACCTTGA	1048
Db	1297	AGTATCAGAGGATGGGAGTGATGTCAGTCCCAACAGAGAAATCTTTTACCT	1356
QY	1049	ATTT 1052	
Db	1357	ACTT 1360	
RESULT 15			
US-10-843-641A-4854			
; Sequence 4854, Application US/10843641A			
; Publication No. US20050064454A1			
; GENERAL INFORMATION:			
; APPLICANT: Avalon Pharmaceuticals, Inc.			
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using			
; FILE REFERENCE: Signature Gene Sets			
; CURRENT APPLICATION NUMBER: US/10/843,641A			
; PRIORITY FILING DATE: 2004-05-12			
; PRIORITY APPLICATION NUMBER: US/09/873,367			
; PRIORITY FILING DATE: 2001-06-05			
; PRIORITY APPLICATION NUMBER: US/09/954,531			
; PRIORITY FILING DATE: 2001-09-18			
; PRIORITY APPLICATION NUMBER: US/09/954,456			
; PRIORITY FILING DATE: 2001-09-25			
; PRIORITY APPLICATION NUMBER: US/09/962,436			
; PRIORITY FILING DATE: 2001-09-25			
; PRIORITY APPLICATION NUMBER: US/09/962,832			
; PRIORITY FILING DATE: 2001-09-25			
; PRIORITY APPLICATION NUMBER: US/09/964,824			
; PRIORITY FILING DATE: 2001-09-27			
; PRIORITY APPLICATION NUMBER: US/09/967,768			
; PRIORITY FILING DATE: 2001-09-28			
; PRIORITY APPLICATION NUMBER: US/09/968,007			
; PRIORITY FILING DATE: 2001-10-02			
; PRIORITY APPLICATION NUMBER: US/09/969,347			
; PRIORITY FILING DATE: 2001-10-02			
; PRIORITY APPLICATION NUMBER: US/09/969,708			
; PRIORITY FILING DATE: 2001-10-03			
; Remaining Prior Application data removed - See File Wrapper or PALM.			
; NUMBER OF SEQ ID NOS: 8447			
; SOFTWARE: PatentIn version 3.0			
; SEQ ID NO 4854			
; LENGTH: 2500			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-843-641A-4854			
Query Match 55.1%; Score 580; DB 22; Length 2500;			
Best Local Similarity 75.1%; Pred. No. 4.6e-157;			
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;			
QY	89	ATGAAGAGCCACTCCTAAGAAAGATTTCTCGCGGGTTTTGTCTATCTTTCCAAATCCAGTACC	148
Db	397	AGGATGAGCCGCTGCTGAGAGAAACCCCGCGGCTTTGTCTATCTTCCCATCGAGTACC	456

Search completed: October 30, 2005, 06:24:12
Job time : 679.5 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	1051.4	99.8	3292	3	CR617553	CR617553	full-length
2	1051.4	99.8	4748	3	HSN802033	AL137348	Homo sapi
3	1039.4	98.7	4650	3	EC042848	BC042948	Homo sapi
4	1009	95.8	3501	3	CR627376	CR627376	Homo sapi
5	943.4	89.6	1056	9	AY398973	AY398973	Homo sapi
6	928	88.1	1083	1	AL547501	AL547501	AL547501
7	828.2	78.7	1008	9	AY398974	AY398974	Pan trogl
8	791.6	75.2	879	5	BQ41857	BQ41857	AGENCOURT
9	778.6	73.9	1043	9	AY398975	BQ41857	AGENCOURT
10	773.2	69.4	892	7	CN163314	AY398975	Mus muscu
11	726.6	69.0	914	7	CF995079	CN163314	952695 MA
12	710.6	67.5	870	7	CO648380	CF995079	AGENCOURT
13	682.8	64.8	804	6	DB656661	CO648380	ILLUMIGEN
14	653.8	62.1	1061	4	BM801298	DB656661	AGENCOURT
15	653.2	62.0	896	5	BU170979	BM801298	AGENCOURT
16	614.8	58.4	685	5	BX951374	BU170979	AGENCOURT
17	578.4	54.9	1571	3	CR603461	BX951374	DKFzp781E
18	578.4	54.9	1573	3	CR625489	CR603461	full-length
19	578.4	54.9	1582	3	CR608076	CR625489	full-length
20	578.4	54.9	1588	3	CR602054	CR608076	full-length
21	578.4	54.9	1592	3	CR621427	CR602054	full-length
22	578.4	54.9	1600	3	CR604378	CR621427	full-length
23	578.4	54.9	1605	3	CR590959	CR604378	full-length
24	578.4	54.9	1612	3	CR614990	CR590959	full-length

QY	61	GACACCAAGAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGATTCTCGC	120
Db	87	GACACCAAGAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGATTCTCGC	146
QY	121	CGGTTTGTCACTTTCCCAATCAGTACCTCGATATTGGAAATGTATAACAGGCACAG	180
Db	147	CGGTTTGTCACTTTCCCAATCAGTACCTCGATATTGGAAATGTATAACAGGCACAG	206
QY	181	GCTTCTTCTGACAGCAGAGGTTGACTTATCAAGAGTCTCCCTCACTGGAACAG	240
Db	207	GCTTCTTCTGACAGCAGAGGTTGACTTATCAAGAGTCTCCCTCACTGGAACAG	266
QY	241	CTTAAGCAGATGAGAGTACTTCTCTCATCTTAGCTTTTTCAGCCAGTGT	300
Db	267	CTTAAGCAGATGAGAGTACTTCTCTCATCTTAGCTTTTTCAGCCAGTGT	326
QY	301	GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGT	360
Db	327	GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGT	386
QY	361	CGCTTTTCTATGGCTTCAATTTCTCATCGAATTTTCACTCAGAGATGTACAGTTTG	420
Db	387	CGCTTTTCTATGGCTTCAATTTCTCATCGAATTTTCACTCAGAGATGTACAGTTTG	446
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATCAATTGAA	480
Db	447	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATCAATTGAA	506
QY	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCGCTTCGATGATAGATAGAAAA	540
Db	507	ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCGCTTCGATGATAGATAGAAAA	566
QY	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTCTTCTCTCAGGA	600
Db	567	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTCTTCTCTCAGGA	626
QY	601	TCTTTTGTCTGTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCATTTTTC	660
Db	627	TCTTTTGTCTGTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCATTTTTC	686
QY	661	AATGAACATCATCAGAGATGAAGACTTCTACTGTGACTTTGCTTGGCTGATGTTCCAA	720
Db	687	AATGAACATCATCAGAGATGAAGACTTCTACTGTGACTTTGCTTGGCTGATGTTCCAA	746
QY	721	TACTTTAGTAAATAGCTTTTCAAGAAAGAGGTCAGGAGATCAITTTGATGCTGTCAA	780
Db	747	TACTTTAGTAAATAGCTTTTCAAGAAAGAGGTCAGGAGATCAITTTGATGCTGTCAA	806
QY	781	ATTGAGCAGGAGTTTAAACAGAGCTTGCAGTTGGCTCATTTGGAAATGAATTGCAAT	840
Db	807	ATTGAGCAGGAGTTTAAACAGAGCTTGCAGTTGGCTCATTTGGAAATGAATTGCAAT	866
QY	841	TTGATGAACACGTACATTCAGTTTGTAGCTGACAGATTACTTGTGCACTTGCATTCTCA	900
Db	867	TTGATGAACACGTACATTCAGTTTGTAGCTGACAGATTACTTGTGCACTTGCATTCTCA	926
QY	901	AAAGGTTTTTTCAGGCAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGAAAA	960
Db	927	AAAGGTTTTTTCAGGCAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGAAAA	986
QY	961	ACAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTTATGCGAGAAC	1020
Db	987	ACAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTTATGCGAGAAC	1046
QY	1021	ACAGATAAGCTTCTTACCTTGCATGCAGATTTT	1053
Db	1047	ACAGATAAGCTTCTTACCTTGCATGCAGATTTT	1079
RESULT 2			
HSM802033			
LOCUS			
HSM802033			
4748 bp			
mRNA			
linear			
HTC 22-SEP-2004			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
CONSRM			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
1. .4748			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="R2PD:DKFZp761E1312"			
/db_xref="taxon:9606"			
/clone="DKFZp761E1312"			
/tissue type="amygdala"			
/clone_lib="761 (synonym: hamy2) . Vector pSport1; host			
DH10B; sites NotI + SalI"			
/dev_stage="adult"			
/notes="ribonucleotide reductase"			
1. .4748			
/genes="DKFZp761E1312"			
47. .1102			
/gene="DKFZp761E1312"			
/codon_start=1			
/product="hypothetical protein"			
/protein_id="CAB70703.2"			
/db_xref="GI:52545612"			
/db_xref="GOA:Q9NTD8"			
/db_xref="UniProt/trEMBL:Q9NTD8"			
/translations="MGDPPEAAGLDQDERSSDTNESEIKNEEPLLRKSSRRFVI			
FPIQYDPIWMYKQQAQAFWTAEEVDLSKDLPHNKLKADKPYFISHILAPFAASDGI			
VNNELVERFSEQVPEARCFYGFQILINHVHSEMSLLIDTIRDPKRFELFNAIE			
TNPYVKKADWLRWIAADRSKTFGRVAFVAAVEGVFSGKSAFAPLWKKRGLMPGLT			
FSNELSRDEGLHCDFACLMFOVLNKPSEERVREIIVDAVKIEQBLTALPVGLIG			
MNCILMKQYIEFVADRLVLELGFVKVQAEPPDFMENISLEKTNFEEKRVSEYQRF			
AWAETTDNDVFTLDADF"			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 1052; Conservative			
99.8%; Score 1051.4; DB 3; Length 4748;			
99.9%; Pred. No. 1.6e-268;			
0; Mismatches 1; Indels 0; Gaps 0;			
QY	1	ATGGCGCACCCGAAAGCGCGGCGCGGCTGGATCAGGATGAGAGATCATCTTCA	60
Db	47	ATGGCGCACCCGAAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA	106
QY	61	GACACCAAGAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGATTCTCGC	120
Db	107	GACACCAAGAGTGAATTAAGTCAATTAAGAGCCACTCTCTTAAGAAAGATTCTCGC	166
QY	121	CGGTTTGTCACTTTTCCAAATCCAGTACCTTGATTTGGAAATGTATAACAGGCACAG	180
Db	167	CGGTTTGTCACTTTTCCAAATCCAGTACCTTGATTTGGAAATGTATAACAGGCACAG	226

QY	181	GCTTCCTTCTGACAGAGAGAGGTTGACTTATCAAAAGGATCTCCCTCACTGGAACAAG	240
Db	227	GCTTCCTTCTGACAGAGAGAGGTCGACTTATCAAGAGATCTCCCTCACTGGAACAAG	286
QY	241	CTTAAAGCAGATGAGAAGTACTTCATCTCTCACTCTTACGCTTTTTCAGCCAGTGTGAT	300
Db	287	CTTAAAGCAGATGAGAAGTACTTCATCTCTCACTCTTACGCTTTTTCAGCCAGTGTGAT	346
QY	301	GGAAATGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT	360
Db	347	GGAAATGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT	406
QY	361	CGCTGTTTCTATGGCTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG	420
Db	407	CGCTGTTTCTATGGCTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG	466
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATCAATTTGAA	480
Db	467	CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATCAATTTGAA	526
QY	481	ACCATGCCCTATGTTAAGAAAACAGATGGGCTTCGCGATGATAGCAGATAGAAA	540
Db	527	ACCATGCCCTATGTTAAGAAAACAGATGGGCTTCGCGATGATAGCAGATAGAAA	586
QY	541	TCTACTTTTGGGAAAGAGTGGTCCCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA	600
Db	587	TCTACTTTTGGGAAAGAGTGGTCCCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA	646
QY	601	TCCTTTGCTGCTATATCTTGGCTAAGAGAGAGGTCTTATCCAGGACTCACTTTTCC	660
Db	647	TCCTTTGCTGCTATATCTTGGCTAAGAGAGAGGTCTTATCCAGGACTCACTTTTCC	706
QY	661	AATGAACCTATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTGTGATGTTTCCAA	720
Db	707	AATGAACCTATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTGTGATGTTTCCAA	766
QY	721	TACTTAGTAAATAGCCCTTCAGAAAGAGGTTCAGGAGATCATTTGTGATGCTGTCAA	780
Db	767	TACTTAGTAAATAGCCCTTCAGAAAGAGGTTCAGGAGATCATTTGTGATGCTGTCAA	826
QY	781	ATTGAGCAGGAGTTTTAAACAGAGCCCTTGCCAGTTGGCCTCATTTGGAATGAATTCATT	840
Db	827	ATTGAGCAGGAGTTTTAAACAGAGCCCTTGCCAGTTGGCCTCATTTGGAATGAATTCATT	886
QY	841	TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGTGGAACTTTGGATTCTCA	900
Db	887	TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGTGGAACTTTGGATTCTCA	946
QY	901	AAGTTTTTTCAGGCAGAAAATCCCTTTGATTTTATGGAACATTTCTTTAGAAGGAAA	960
Db	947	AAGTTTTTTCAGGCAGAAAATCCCTTTGATTTTATGGAACATTTCTTTAGAAGGAAA	1006
QY	961	ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGTTATGSCAGAAACC	1020
Db	1007	ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGTTATGSCAGAAACC	1066
QY	1021	ACAGATAACGCTTTCACCTTGGATGCAGATTTT	1053
Db	1067	ACAGATAACGCTTTCACCTTGGATGCAGATTTT	1099
RESULT 3	BC042948	Homo sapiens cdna clone IMAGE:4798175, containing frame-shift errors.	
LOCUS	BC042948	4650 bp mRNA linear HTC 12-OCT-2004	
DEFINITION	BC042948	Homo sapiens (human)	
ACCESSION	BC042948		
VERSION	BC042948.2	GI:34194000	
KEYWORDS	HTC		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			

Db 72 ATGGCGCACCAGAAAGCCGCGCGCTGGATCAGGATGAGAGATCATCTTCA 131
Qy 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCCTTAAGAAAGAGTTCTCGC 120
Db 132 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAG-CACTCCTAAGAAAGAGTTCTCGC 190
Qy 121 CGGTTTGTCATCTTTCCAAATCAGTACCTGATATTTGGAAAAATGTATAACAGGCACAG 180
Db 191 CGGTTTGTCATCTTTCCAAATCAGTACCTGATATTTGGAAAAATGTATAACAGGCACAG 250
Qy 181 GCTTCCTTCTGGACAGCAAGAGAGTTGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 240
Db 251 GCTTCCTTCTGGACAGCAAGAGAGTGACATTATCAAAAGGATCTCCCTCACTGGAAACAAG 310
Qy 241 CTTAAGCAGATGAGAGTACTTCACTCTCACATCTTTAGCCTTTTTTGCAGCCAGTGAT 300
Db 311 CTTAAGCAGATGAGAGTACTTCACTCTCACATCTTTAGCCTTTTTTGCAGCCAGTGAT 370
Qy 301 GGAATTGTAAATGAAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGTTCAGAGGCT 360
Db 371 GGAATTGTAAATGAAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGTTCAGAGGCT 430
Qy 361 CGCTGTTTCTATGGCTTTCAAATCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db 431 CGCTGTTTCTATGGCTTTCAAATCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 490
Qy 421 CTGATAGACATTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAAATGCAATTGAA 480
Db 491 CTGATAGACATTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAAATGCAATTGAA 550
Qy 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCTTTGCGATGATAGCAGATAGAAAA 540
Db 551 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCTTTGCGATGATAGCAGATAGAAAA 610
Qy 541 TCTACTTTTGGGAAAAGAGTGGTGCCCTTTGCTGCTGTAGAAGGAGTTTCTTCTCAGGA 600
Db 611 TCTACTTTTGGGAAAAGAGTGGTGCCCTTTGCTGCTGTAGAAGGAGTTTCTTCTCAGGA 670
Qy 601 TCTTTTGGCTGCTATATCTTGGCTAAGAAAGAGAGTCTTTATGCCAGGACTCACTTTTCC 660
Db 671 TCTTTTGGCTGCTATATCTTGGCTAAGAAAGAGAGTCTTTATGCCAGGACTCACTTTTCC 730
Qy 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCTGATGTTCCAA 720
Db 731 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCTGATGTTCCAA 790
Qy 721 TACTTAGTAAATAAGCCTTTCAGAAAAAGGGTCAAGGAGATCATTTGTTGATGCTGTCAA 780
Db 791 TACTTAGTAAATAAGCCTTTCAGAAAAAGGGTCAAGGAGATCATTTGTTGATGCTGTCAA 850
Qy 781 ATTGAGCAGGAGTTTAAACGAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTTGCATT 840
Db 851 ATTGAGCAGGAGTTTAAACGAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTTGCATT 910
Qy 841 TTGATGAACAGTACATTGAGTTTCTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Db 911 TTGATGAACAGTACATTGAGTTTCTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 970
Qy 901 AAGGTTTTTTCAGGCAGAAAAATCCTTTTGATTTTATGGAACCAATTTCTTTAGAAGAAAA 960
Db 971 AAGGTTTTTTCAGGCAGAAAAATCCTTTTGATTTTATGGAACCAATTTCTTTAGAAGAAAA 1030
Qy 961 ACAAAATTTCTTGAGAAAACGAGTTTTCAGAGTATCAGCGTTTTTTCAGTTATGGCAGAAACC 1020
Db 1031 ACAAAATTTCTTGAGAAAACGAGTTTTCAGAGTATCAGCGTTTTTTCAGTTATGGCAGAAACC 1090
Qy 1021 ACAGATAACGCTTTCACCTTGGATGCAATTTT 1053
Db 1091 ACAGATAACGCTTTCACCTTGGATGCAATTTT 1123

LOCUS CR627376 3501 bp mRNA linear HTC 22-SEP-2004
DEFINITION Homo sapiens mRNA; cDNA DKFZp686M05248 (from clone DKFZp686M05248).
ACCESSION CR627376
VERSION CR627376.1 GI:50949847
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3501)
AUTHORS Ottenwaelder,B., Obermaier,B., Deutschenbaur,S., Schaipp,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
CONSTRM The German cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix (Martinsried/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp686M05248) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp686M05248
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.
FEATURES
Location/Qualifiers
1..3501
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686M05248"
/tissue_type="salivary gland"
/clone_lib="686 (synonym: hlcc3) . Vector pSport1_sfi; host DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
/note="ribonucleotide reductase M2 B (TP53 inducible)"
1..3501
/gene="DKFZp686M05248"
332..1225
/gene="DKFZp686M05248"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAHI0473.1"
/db_xref="GI:50949848"
/translation="MYKQAQSFWTAEVDLSKDLPHWNKLKADKVFYFISHLAFFAA SDGIYVENLVERFVSQVPEARCFYGFOLIENVHSEMYSLIDTVIRDPKKRFLF NAIETMPYVKKADWALRWIADRKSTFGERVVAFAAVEGVFPSSFAAI FMLKKRGLM PGLTFSNELISRDEGLHCDFACLMFQYLVNKPPEERVREIIVDAVKLEQSFLEALPV GLIGMNCIMKQYIEFVADRLVLVLFSGVSKVQAEVNFDFMENISLGGKTNFFEKRVSE YORFVMAETTDNVFTLDADF"
ORIGIN
Query Match 95.8%; Score 1009; DB 3; Length 3501;
Best Local Similarity 98.5%; Pred. No. 2.9e-257;
Matches 1018; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
Qy 21 GGAAGCGCGCGGCTGGATCAGGATGAGAGATCATCTTCAGACACCAACGAAAGTGAAT 80
Db 190 GGACGACGCTTGGAGCGCTCAGCCGGGAGATCATCTTCAGACACCAACGAAAGTGAAT 249
Qy 81 AAAGTCAAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGCCGGTTGTCTATCTTTCAAT 140
Db 250 AAAGTCAAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGCCGGTTGTCTATCTTTCAAT 309
Qy 141 CCAGTACCCCTGATATTGGAAAAATGTATAACAGGCACAGGCTTCCTCTGGACAGCAGA 200
Db 310 CCAGTACCCCTGATATTGGAAAAATGTATAACAGGCACAGGCTTCCTCTGGACAGCAGA 369
Qy 201 AGAGGTTGACTTATCAAGAGATCTCCCTCACTCGAAACAAGCTTTAAAGCAGATGAGAAAGTA 260
Db 370 AGAGGTCGACTTATCAAGAGATCTCCCTCACTCGAAACAAGCTTTAAAGCAGATGAGAAAGTA 429

```
QY 261 CTTTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGTGTGGAATTTGTAATGAAATTTT 320
Db 430 CTTTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGTGTGGAATTTGTAATGAAATTTT 489
QY 321 GGTGGAGCGCTTTAGTCAGAGGTGAGGTTTCAGAGGCTCGCTGTTTCTATGGCTTCA 380
Db 490 GGTGGAGCGCTTTAGTCAGAGGTGAGGTTTCAGAGGCTCGCTGTTTCTATGGCTTCA 549
QY 381 AATTCTCATCGAAGTGTTCACATCAGAGATGTACAGTTTCTGTATAGACATTTACATCAG 440
Db 550 AATTCTCATCGAAGTGTTCACATCAGAGATGTACAGTTTCTGTATAGACATTTACATCAG 609
QY 441 AGATCCCAAGAAAGGAAATTTTATTTAATGCAATTTGAAACCATGCGCTATGTTAAGAA 500
Db 610 AGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGAAACCATGCGCTATGTTAAGAA 669
QY 501 AAAAGCAGATTGGGCTTCGGATGGATAGCAGATAGAAAATCTACTTTTGGGAAAGAGT 560
Db 670 AAAAGCAGATTGGGCTTCGGATGGATAGCAGATAGAAAATCTACTTTTGGGAAAGAGT 729
QY 561 GGTGGCCTTTGCTGTAGAGGAGTTTCTTCTCAGGATCTTTTCTGCTCTATATCTG 620
Db 730 GGTGGCCTTTGCTGTAGAGGAGTTTCTTCTCAGGATCTTTTCTGCTCTATATCTG 789
QY 621 GCTAAAGAGAGAGGCTTTATGCGCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGA 680
Db 790 GCTAAAGAGAGAGGCTTTATGCGCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGA 849
QY 681 TGAAGGACTTCACTGTGACTTTTCTGCTGTAGTGTCCAAATCTAGTAAATAGGCTTC 740
Db 850 TGAAGGACTTCACTGTGACTTTTCTGCTGTAGTGTCCAAATCTAGTAAATAGGCTTC 909
QY 741 AGAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTTAGCAGAGTTTAAAC 800
Db 910 AGAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTTAGCAGAGTTTAAAC 969
QY 801 AGAAGCCTTGGCAGTTGGCTCATTTGGAATGAATTCGATTTGATGAACAGTACATTGA 860
Db 970 AGAAGCCTTGGCAGTTGGCTCATTTGGAATGAATTCGATTTGATGAACAGTACATTGA 1029
QY 861 GTTTGTAGCTGACAGATTACTTGTGGAACTTGGAATCTCAAGGTTTTCAGGACGAGAAA 920
Db 1030 GTTTGTAGCTGACAGATTACTTGTGGAACTTGGAATCTCAAGGTTTTCAGGACGAGAAA 1089
QY 921 TCTTTTGTATTTATGGAACAACTTTCTTTAGAGGAAACAAATTTCTTTTGGAGAAACG 980
Db 1090 TCTTTTGTATTTATGGAACAACTTTCTTTAGAGGAAACAAATTTCTTTTGGAGAAACG 1149
QY 981 AGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACACAGATAACGTTTACCTT 1040
Db 1150 AGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACACAGATAACGTTTACCTT 1209
QY 1041 GGATGCAGATTTT 1053
Db 1210 GGATGCAGATTTT 1222

RESULT 5
LOCUS AY398973
DEFINITION Homo sapiens HCM0069 gene, VIRUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY398973
VERSION AY398973.1 GI:39754962
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1056)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
```

```
TITLE
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1056)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferred nonneutral evolution from human-chimp-mouse orthologous
gene trios
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source 1..1056
location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1056
/locus_tag="HCM0069"
gene
ORIGIN
Query Match 89.6%; Score 943.4; DB 9; Length 1056;
Best Local Similarity 89.6%; Pred. No. 6.2e-240;
Matches 944; Conservative 0; Mismatches 109; Indels 0; Gaps 0;
```

```
QY 1 ATGGGCGACCCGGAAAGCGCGGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGGCGACCCGGAAAGCGCGGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCG 120
Db 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCG 120
QY 121 CGGTTTGTCTATCTTCCAATCCAGTACCTGATATTTGGAAAAATGATAAAGAGCAG 180
Db 121 CGGTTTGTCTATCTTCCAATCCAGTACCTGATATTTGGAAAAATGATAAAGAGCAG 180
QY 181 GCTTCTCTTCGGACAGCAGAGAGGTTGACTTATCAAGAGGATCTCCCTCAGTGAACAG 240
Db 181 GCTTCTCTTCGGACAGCAGAGAGGTTGACTTATCAAGAGGATCTCCCTCAGTGAACAG 240
QY 241 CTTAAAGCAGATGAGAAAGTACTTTCATCTCTCACTCTTAGCCTTTTTCGAGCCAGTGT 300
Db 241 CTTAAAGCAGATGAGAAAGTACTTTCATCTCTCACTCTTAGCCTTTTTCGAGCCAGTGT 300
QY 301 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGGTTCCAGAG 360
Db 301 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGGTTCCAGAG 360
QY 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGATGTTCTACTCAGAGATGTACAG 420
Db 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGATGTTCTACTCAGAGATGTACAG 420
QY 421 CTGATGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTTGAA 480
Db 421 CTGATGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCGCTTGGCATGGATAGCAGATAGAAA 540
Db 481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCCTTTTGTCTGTAGAGAGATTTTCTTCTCAGGA 600
Db 541 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 600
QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAAAGAGGTTCTATGCGCAGGAGTCACTTTTTC 660
Db 601 TCTTTTGTCTATATTTCTGGCTAAAGAAAGAGGTTCTATGCGCAGGAGTCACTTTTTC 660
```

Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferred nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 1056)
Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.

source 1..1056
location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1056
/locus_tag="HCM0069"

Query Match 89.6%; Score 943.4; DB 9; Length 1056;
Best Local Similarity 89.6%; Pred. No. 6.2e-240;
Matches 944; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

```
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db |||||
QY 721 TACTTAGTAATAAGCCTTCAGAAGAAAGGTCAGGAGATCAATGTTGATGCTGTCAAA 780
Db |||||
QY 781 ATTGACACGAGGTTTAAACAGAGAGCCTTGCCAGTTGGCCTCATTCGGAATGCAATTCGANT 840
Db |||||
QY 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGAATTCCTCA 900
Db |||||
QY 901 AAGGTTTTTCAGGACAGAAATCCTTTTGTATTTATGGAACAAATTTCTTTAGAGGAA 960
Db |||||
QY 961 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGGTTTTCAGTATGCGAGAAACC 1020
Db |||||
QY 1021 ACAGATAAGCTCTTCACTTGGATGAGATTTT 1053
Db |||||
QY 1021 ACAGATAAGCTCTTCACTTGGATGAGATTTT 1053
Db |||||
```

RESULT 6

```
AL547501
LOCUS 1083 bp mRNA linear EST 25-MAR-2004
DEFINITION AL547501 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1011YF14 5-PRIME, mRNA sequence.
ACCESSION AL547501
VERSION AL547501.3 GI:45747948
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001).
On Feb 15, 2001 this sequence version replaced gi:31269332.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4436.r
```

For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?s=CS0D1011DC070P1&c=4436.r>.

FEATURES

```
source
1. .1083
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /clone="CS0D1011YF14"
   /tissue_type="PLACENTA COT 25-NORMALIZED"
   /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
   /note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

Query Match

88.1%; Score 928; DB 1; Length 1083;

RESULT 7

```
Best Local Similarity 97.9%; Pred. No. 7.9e-236;
Matches 975; Conservative 6; Mismatches 11; Indels 4; Gaps 4;
QY 1 ATGGGCGACCCGAAAGCGCGGCGG-CCTGGATCAGATCAGAGATCATCTTC 59
Db |||||
QY 27 ATGGGCGACCCGAAAGCGCGGCGGCGGCGGCTGGATCAGATCAGAGATCATCTTC 86
Db |||||
QY 60 AGACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTCTCG 119
Db |||||
QY 87 AGACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTCTCG 146
Db |||||
QY 120 CCGGTTTGTTCATCTTCCAAATCCAGTACCCTGATATTTGGAATAATGATATAACAGGCACA 179
Db |||||
QY 147 CCGGTTTGTTCATCTTCCAAATCCAGTACCCTGATATTTGGAATAATGATATAACAGGCACA 206
Db |||||
QY 180 GGCCTTCCTCTGACAGCAGAGAGGTCGACTTATCAAGAGATCTCCCTCACTGGAACAA 239
Db |||||
QY 207 GGCCTTCCTCTGACAGCAGAGAGGTCGACTTATCAAGAGATCTCCCTCACTGGAACAA 266
Db |||||
QY 240 GCTTAAAGCAGATGAGAAAGTACTTCACTCTCACTCTAGCCCTTTTTCAGCCAGTGA 299
Db |||||
QY 267 GCTTAAAGCAGATGAGAAAGTACTTCACTCTCACTCTAGCCCTTTTTCAGCCAGTGA 326
Db |||||
QY 300 TGAATTTGTAATGAAAAATTTGCTGGAGCGCTTTAGTCAGAGGTCAGGTTCCAGAGGC 359
Db |||||
QY 327 TGAATTTGTAATGAAAAATTTGCTGGAGCGCTTTAGTCAGAGGTCAGGTTCCAGAGGC 386
Db |||||
QY 360 TCCTGTTTCTATGGCTTTCAAAATTCATCTCGAGAATGTTCACTCAGAGATGTCAGTTT 419
Db |||||
QY 387 TCCTGTTTCTATGGCTTTCAAAATTCATCTCGAGAATGTTCACTCAGAGATGTCAGTTT 446
Db |||||
QY 420 GCTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGA 479
Db |||||
QY 447 GCTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGA 506
Db |||||
QY 480 AACCATGCCCTATGTTTAAAGAAAAAGCAGATTTGGGCTTGGATGGAATAGAGATAGAAA 539
Db |||||
QY 507 AACCATGCCCTATGTTTAAAGAAAAAGCAGATTTGGGCTTGGATGGAATAGAGATAGAAA 566
Db |||||
QY 540 ATCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGAGTTTCTTCTCAGG 599
Db |||||
QY 567 ATCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGAGTTTCTTCTCAGG 626
Db |||||
QY 600 ATCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGTCCTATATGCCAGGACTCAGTTTTC 659
Db |||||
QY 627 ATCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGTCCTATATGCCAGGACTCAGTTTTC 686
Db |||||
QY 660 CAATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTTCTGCTGCTGATGTTCCA 719
Db |||||
QY 687 CAATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTTCTGCTGCTGATGTTCCA 746
Db |||||
QY 720 ATACTTAGTAATAAGCCTTCAGAAGAAAGGTCAGGAGATCAATGTTGATGCTGTCAAA 779
Db |||||
QY 747 ATACTTAGTAATAAGCCTTCAGAAGAAAGGTCAGGAGATCAATGTTGATGCTGTCAAA 806
Db |||||
QY 780 AATTGACAGAGCTTTTAAACAGAGCCTTGCAGTTGGCCTCATTCGAATGAATTCGAT 839
Db |||||
QY 807 AATTGACAGAGCTTTTAAACAGAGCCTTGCAGTTGGCCTCATTCGAATGAATTCGAT 866
Db |||||
QY 840 TTTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTAATCTG-TGGAACCTTGGAATTC 898
Db |||||
QY 867 TTTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTAATCTG-TGGAACCTTGGAATTC 926
Db |||||
QY 899 CAAAGGTTTTTCAGGCAAGAAATCCTTTTGAATTTTATGGAATAATTTCTTTTAGAAGAA 958
Db |||||
QY 927 CAAAGGTTTTTCAGGCAAGAAATCCTTTTGAATTTTATGGAATAATTTCTTTTAGAAG-A 984
Db |||||
QY 959 AACAATAATTTCTTTGAGAAACGAGTTTCAGAGATTC 994
Db |||||
QY 985 AAAAAAATTTCTTTTGAACACGAGTTTMRAGMWCAGC 1020
Db |||||
```

```
AY398974      1008 bp      DNA      linear      GSS 12-DEC-2003
LOCUS      Pan troglodytes HCM0069 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION      genomic survey sequence.
ACCESSION      AY398974
VERSION      AY398974.1 GI:39754963
KEYWORDS      GSS.
SOURCE      Pan troglodytes (chimpanzee)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
REFERENCE      1 (bases 1 to 1008)
AUTHORS      Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
      Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
      Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
      Adams,M.D. and Cargill,M.
TITLE      Direct Submission
JOURNAL      Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
COMMENT      Rockville, MD 20850, USA
      This sequence was made by sequencing genomic exons and ordering
      them based on alignment.
FEATURES             Location/Qualifiers
     source           1..1008
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     <1..>1008
                     /locus_tag="HCM0069"
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
```

```
Query Match      78.7%; Score 828.2; DB 9; Length 1008;
Best Local Similarity 82.6%; Pred. No. 3e-209;
Matches 830; Conservative 0; Mismatches 175; Indels 0; Gaps 0;

QY 49 AGATCATCTTCAGACACCAACGAAAGTGAATTAAGTCAAAAGTGAAGAGCCACTCTCTAAGA 108
Db 1 AGATCATCTTCAGACACCAACGAAAGTGAATTAAGTCAAAAGTGAAGAGCCACTCTCTAAGA 60

QY 109 AAGAGTTCTCGCGGTTTGTCTCTCTTCCAAATCCAGTACCCTGATATTTGGAAATGTAT 168
Db 61 AAGAGTTCTCGCGGTTTGTCTCTCTTCCAAATCCAGTACCCTGATATTTGGAAATGTAT 120

QY 169 AAACAGGCACAGGCTCTCTTCTGGACAGAGAGGTTGACTTATCAAAAGGATCTCCCT 228
Db 121 AAACAGGCACAGGCTCTCTTCTGGACAGAGAGGTTGACTTATCAAAAGGATCTCCCT 180

QY 229 CACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTAGCCTTTTT 288
Db 181 CACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTAGCCTTTTT 240

QY 289 GCAGCAGTGTGGAAATTTGAATTAAGTGAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAG 348
Db 241 GCAGCAGTGTGGAAATTTGAATTAAGTGAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAG 300

QY 349 GTTCAGAGGCTCGCTGTTTCTATGCTTTTCAAAATTTCTCATCGAGAATTTCACTCAGAG 408
Db 301 GTTCAGAGGCTCGCTGTTTCTATGCTTTTCAAAATTTCTCATCGAGAATTTCACTCAGAG 360

QY 409 ATGTACAGTTTCTCATGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTT 468
Db 361 ATGTACAGTTTCTCATGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTT 420

QY 469 AATGCAATTGAACCAACATGCCCTATGTTAAGAAAGAGCAGATTTGGCCTTTGCATCGATA 528
```

```
421 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 480
QY 529 GCAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCCCTTGGCTGCTGTAAGAGGAGTT 588
Db 481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540
QY 589 TTCTTCTCAGGATCTTTTGTCTGTATATTTCTGGCTAAAGAGAGAGGCTTATATGCCAGGA 648
Db 541 TTCTTCTCAGGATCTTTTGTCTGTATATTTCTGGCTAAAGAGAGAGGCTTATATGCCAGGA 600
QY 649 CTCACTTTTCCAAATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGC 708
Db 601 NTCACATTTTCCAAATGAACATCANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 660
QY 709 CTGATGTTTCCAAATGACTTAGTAAATAGCCCTTCAGAAAGAGGTCAGGAGATCATTTGT 768
Db 661 CTGATGTTTCCAAATGACTTAGTAAATAGCCCTTCAGAAAGAGGTCAGGAGATCATTTGT 720
QY 769 GATGCTGTCAAAATTTGAGCAGGAGGATTTTAAACAGAGCCTTGCCAGTTGGCCTCATTTGA 828
Db 721 GATGCTGTCAAAATTTGAGCAGGAGGATTTTAAACAGAGCCTTGCCAGTTGGCCTCATTTGA 780
QY 829 ATGAATTTGCAATTTTGAATGAACAGTACATTTGAGTTTGTAGCTGACAGATTTCTTTGGA 888
Db 781 ATGAATTTGCAATTTTGAATGAACAGTACATTTGAGTTTGTAGCTGACAGATTTCTTTGGA 840
QY 889 CTTGGATTTCAAAAGGTTTTCAGGCAGAGAAATCTTTGATTTATTTGGAAGAAATTTCT 948
Db 841 CTTGGATTTCAAAAGGTTTTCAGGCAGAGAAATCTTTGATTTATTTGGAAGAAATTTCT 900
QY 949 TTGAAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCATCAGCTTTTGCAGTT 1008
Db 901 TTGAAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCATCAGCTTTTGCAGTT 960
QY 1009 ATGCAGAAACCCACAGATACGCTTTCACCTTGGATGACAGATTTT 1053
Db 961 ATGCAGAAACCCACAGATACGCTTTCACCTTGGATGACAGATTTT 1005

RESULT 8
LOCUS      BQ441857
DEFINITION      BQ441857
ACCESSION      BQ441857
VERSION      BQ441857.1 GI:21180933
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 879)
AUTHORS      NIH-MGC http://mgc.ncl.nih.gov/
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
      Email: cgabbs@mail.nih.gov
      Tissue Procurement: ATCC
      cDNA Library Preparation: Life Technologies, Inc.
      DNA Sequencing by: Agencourt Bioscience Corporation
      Clone distribution: MGC clone distribution information can be
      found through the I.M.A.G.E. Consortium/LLNL at:
      http://image.llnl.gov
      Plate: LLAM13208 row: d column: 20
      High quality sequence stop: 662.
      Location/Qualifiers
      1..879
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:6014179"
      /tissue_type="embryonal carcinoma, cell line"
FEATURES             source
```



```

/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 92"
/notes="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH MGC library."

```

/lab_host="DH10B (phage-resistant)" /clone_lib="NIH_MGC_92" /note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."			
ORIGIN	Query Match Best Local Similarity 98.6%; Score 791.6; DB 5; Length 879; Matches 820; Conservative 0; Mismatches 9; Indels 3; Gaps 2;		
Qy	1 ATGGCGACCCGGAAAGCGCGGAGCGCGCGCTGGATCAGATGAGAGATCATCTTCA 60 		
Db	22 ATGGCGACCCGGAAAGCGCGGAGCGCGCGCTGGATCAGATGAGAGATCATCTTCA 81 		
Qy	61 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCATCTCTAAGAAAGATTCTCGC 120 		
Db	82 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCATCTCTAAGAAAGATTCTCGC 141 		
Qy	121 CGGTTTCTCATCTTTCCAAATCCAGTACCCCTGATATTTGGAAATGTATAACAGGCACAG 180 		
Db	142 CGGTTTGTCTTTTCCAAATCCAGTACCCCTGATATTTGGAAATGTATAACAGGCACAG 201 		
Qy	181 GCTTTCCTTTCTGGACAGCAGAGAGGTTGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 240 		
Db	202 GCTTTCCTTTCTGGACAGCAGAGAGGTCGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 261 		
Qy	241 CTTAAAGCAGATGAGAAGTACTTCATCTCTCACATCTAGCCCTTTTTCGAGCGCAGTGAT 300 		
Db	262 CTTAAAGCAGATGAGAAGTACTTCATCTCTCACATCTAGCCCTTTTTCGAGCGCAGTGAT 321 		
Qy	301 GGAATTCGAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360 		
Db	322 GGAATTCGAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 381 		
Qy	361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATTTTCACTCAGAGATGTACAGTTTG 420 		
Db	382 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATTTTCACTCAGAGATGTACAGTTTG 441 		
Qy	421 CTGATGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480 		
Db	442 CTGATGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 501 		
Qy	481 ACCATGCCCTATGTTAAGAAAGAGCAGATTGGCCCTTTCGATGGATAGCAGATAGAAAA 540 		
Db	502 ACCATGCCCTATGTTAAGAAAGAGCAGATTGGCCCTTTCGATGGATAGCAGATAGAAAA 561 		
Qy	541 TCTACTTTTGGGAAAGAGTGGTGGCCCTTTGCTGCTGTAGAAGAGATTTTCTTCTCAGGA 600 		
Db	562 TCTACTTTTGGGAAAGAGTGGTGGCCCTTTGCTGCTGTAGAAGAGATTTTCTTCTCAGGA 621 		
Qy	601 TCTTTTGGTGTCTATTTCTGGCTTAAGAGAGAGGTCTTATGCGAGACTCACTTTTTC 660 		
Db	622 TCTTTTGGTGTCTATTTCTGGCTTAAGAGAGAGGTCTTATGCGAGACTCACTTTTTC 681 		
Qy	661 AATGAACTCATCAGCAGATGAGAGCTTCACTGTCATTTGCTTCCTGCTGATGTTCCAA 720 		
Db	682 AATGAACTCATCAGCAGATGAGAGCTTCACTGTCATTTGCTTCCTGCTGATGTTCCAA 741 		
Qy	721 TACTTAGTAAATAAGCCCTTCAGAAAGAGGG--TCAGGGAGATCATTTGTTGATGCTGTCA 778 		
Db	742 TACTTAGTAAATAAGCCCTTCAGAAAGAGGGTTCAGGGAGATCATTTGTTGATGCTGTCA 801 		
Qy	779 AAATT-GAGCAGGAGTTTTTAAACAGAGCCCTTGCAGTTGGCCTCATTTGAA 829 		
Db	802 AAATTGGAGCAGGAGTTTTTAAACAGAGCCCTTGCAGTTGGCCTCAATTGGA 853 		
RESULT 9	AY398975	1043 bp	DNA linear GSS 12-DEC-2003
LOCUS			

RESULT 9
AY398975
LOCUS

QY 212 TATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTTCATCTCTC 271
| | | | |
Db 1 TATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTTCATCTCTC 60
| | | | |
QY 272 ACATCTTAGC- CTTTTCCTCAGCCAGTGTAAATTTGTAATGAAATTTGGTGGAGCGC 330
| | | | |
Db 61 ACATCTTAGCNCCTTTTTCAGCCAGTGTAAATTTGTAATGAAATTTGGTGGAGCGC 120
| | | | |
QY 331 TTATAGTCAGAGGTGCGAGTTCAGAGGCTCGCTGTTCTATGGCTTTCAAAATTCATC 390
| | | | |
Db 121 TTATAGTCAGAGGTGCGAGTTCAGAGGCTCGCTGTTCTATGGCTTTCAAAATTCATC 180
| | | | |
QY 391 GAGAATGTTTCACTCAGAGATGTACAGTTCCTGTATAGACACTTACATCAGAGATCCCAAG 450
| | | | |
Db 181 GAGAATGTTTCACTCAGAGATGTACAGTTCCTGTATAGACACTTACATCAGAGATCCCAAG 240
| | | | |
QY 451 AAAAGGAAATTTTATTTAATGCAATGAAACCATGCCCTATGTTTAAGAAAAAGCAGAT 510
| | | | |
Db 241 AAAAGGAAATTTTATTTAATGCAATGAAACCATGCCCTATGTTTAAGAAAAAGCAGAT 300
| | | | |
QY 511 TGGGCTTCGGATGATAGACAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCCTTT 570
| | | | |
Db 301 TGGGCTTCGGATGATAGACAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCCTTT 360
| | | | |
QY 571 GCTGCTGTAGAAGGAGTTTCTCTCAGGATCTTTTGGCTGTATATTTCTGGCTTAAAGAG 630
| | | | |
Db 361 GCTGCTGTAGAAGGAGTTTCTCTCAGGATCTTTTGGCTGTATATTTCTGGCTTAAAGAG 420
| | | | |
QY 631 AGAGGTCTTTATGCGAGGACTCACTTTTCCAACTGAACATCATCAGCAGAGATGAAGACTT 690
| | | | |
Db 421 AGAGGTCTTTATGCGAGGACTCACTTTTCCAACTGAACATCATCAGCAGAGATGAAGACTT 480
| | | | |
QY 691 CACTGTGACTTTGCTTGCCTGATGTTCCAACTACTTAGTAAATAAGCCTTCAGAAAAAG 750
| | | | |
Db 481 CACTGTGACTTTGCTTGCCTGATGTTCCAACTACTTAGTAAATAAGCCTTCAGAAAAAG 540
| | | | |
QY 751 GTCAGGAGATCATTTGTCATGCTGTCAAAATTGACGAGGAGTTTAAACAGAGCCTTG 810
| | | | |
Db 541 GTCAGGAGATCATTTGTCATGCTGTCAAAATTGACGAGGAGTTTAAACAGAGCCTTG 600
| | | | |
QY 811 CCAGTTGGCCTCATTTGGAATGAATTCATTTTATGAAACAGTACATTCAGTTTGTAGCT 870
| | | | |
Db 601 CCAGTTGGCCTCATTTGGAATGAATTCATTTTATGAAACAGTACATTCAGTTTGTAGCT 660
| | | | |
QY 871 GACAGATTACTGTGGAATCTGGATTCTCAAGGTTTTTCAGCAGAAAAATCCTTTGTAT 930
| | | | |
Db 661 GACAGATTACTGTGGAATCTGGAT--CTCAAGGTTTTTCAGCAGAAAAATCCTTTGTAT 718
| | | | |
QY 931 TTTATGGAACAATTT 946
| | | | |
Db 719 TATGGAACAATTTCTTT 734
| | | | |

RESULT 14
BM801298
LOCUS
DEFINITION
AGENCOURT_6422860 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5586693
5', mRNA sequence.
ACCESSION
BM801298
VERSION
BM801298.1 GI:19118121
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1061)
NIH-MGC <http://mgc.nci.nih.gov/>.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12354 row: p column: 22
High quality sequence stop: 631.
Location/Qualifiers
1..1061
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5586693"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

FEATURES
source

ORIGIN

Query Match 62.1%; Score 653.8; DB 4; Length 1061;
Best Local Similarity 97.4%; Pred.No. 8.9e-163;
Matches 717; Conservative 0; Mismatches 13; Indels 6; Gaps 5;
QY 1 ATGGGCGACCCGGAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
| | | | |
Db 85 ATGGGCGACCCGGAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 144
| | | | |
QY 61 GACACCAAGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
| | | | |
Db 145 GACACCAAGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 204
| | | | |
QY 121 CGTTTTGTCACTTTTCCAACTCCAGTACCCTGATATTTGGAATATGTAACAGGCACAG 180
| | | | |
Db 205 CGTTTTGTCACTTTTCCAACTCCAGTACCCTGATATTTGGAATATGTAACAGGCACAG 264
| | | | |
QY 181 GCTTCCTTCTGGACAGCAGAGAGGTGTACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 240
| | | | |
Db 265 GCTTCCTTCTGGACAGCAGAGAGGTGTACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 324
| | | | |
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTAGCCCTTTTTCAGCCAGCAGTGAT 300
| | | | |
Db 325 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTTAGCCCTTTTTCAGCCAGCAGTGAT 384
| | | | |
QY 301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTTCAGAGGCT 360
| | | | |
Db 385 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTTCAGAGGCT 444
| | | | |
QY 361 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 420
| | | | |
Db 445 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 504
| | | | |
QY 421 CTGATAGACACTTTACATCAGAGATCCCAAGAAAGGGAATTTTTTATTAATGCAATTGAA 480
| | | | |
Db 505 CTGATAGACACTTTACATCAGAGATCCCAAGAAAGGGAATTTTTTATTAATGCAATTGAA 564
| | | | |
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTTCGATGGATAGCAGATAGAAAA 540
| | | | |
Db 565 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTTCGATGGATAGCAGATAGAAAA 624
| | | | |
QY 541 TCTACTTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTAGAAAGAGTTTTTCTTCTCAGGA 600
| | | | |
Db 625 TCTACTTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTAGAAAGAGTTTTTCTTCTCAGGA 684
| | | | |
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTTATATGCGAGGACTCACTTTTTTC 660
| | | | |
Db 685 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTC-TATGCGAGGACTCACTTTTTTTC 743
| | | | |
QY 661 AATGAACACTCAT-CAGCAGAGAT-GAAGGACTTCACCTGTGA--CTTTGCTTGCCTGAT-GT 715
| | | | |

```
Db 744 CATGCACTCATCCACGAGATGGAGGAGCTCCCTGTGAACCTTTGCCCTGGCTGATGT 803
QY 716 TCCAACTACTAGTAA 731
Db 804 TCCAACTACTNAAGAA 819

RESULT 15
LOCUS BUI70979 896 bp mRNA linear EST 04-SEP-2002
DEFINITION AGENCOURT_7940338 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6143754
5', mRNA sequence.
ACCESSION BUI70979
VERSION BUI70979.1 GI:22684963
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 896)
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM13467 row: k column: 19
High quality sequence stop: 619.
FEATURES
Location/Qualifiers
1..896
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6143754"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_67"
/Note="Organ: eye; Vector: pCMV-SPORT6; Site:1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
ORIGIN
Query Match 62.0%; Score 653.2; DB 5; Length 896;
Best Local Similarity 98.5%; Pred. No. 1.2e-162;
Matches 670; Conservative 0; Mismatches 8; Indels 2; Gaps 1;
QY 1 ATGGGCGACCCGGAAGCGCGGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
Db 55 ATGGGCGACCCGGAAGCGCGGCGGCTGGATCAGGATCAGAGATCATCTTCA 114
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 115 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 174
QY 121 CGGTTTGTCACTTTTCCAATCCAGTACCTGTATTTGGAAAATGTPATAAACAGGCACAG 180
Db 175 CGGTTTGTCACTTTTCCAATCCAGTACCTGTATTTGGAAAATGTPATAAACAGGCACAG 234
QY 181 GCTTCCTTCTGGACAGCAAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAAACAAG 240
Db 235 GCTTCCTTCTGGACAGCAAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAAACAAG 294
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCTCATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db 295 CTTAAAGCAGATGAGAAGTACTTCTCTCATCTTAGCCCTTTTTCAGCCAGTGAT 354
```

```
QY 301 GGAATTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
Db 355 GGAATTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 414
QY 361 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 415 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 474
QY 421 CTGATAGACACTTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTATTAATGCAATTGAA 480
Db 475 CTGATAGACACTTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTATTAATGCAATTGAA 534
QY 481 ACCATGCCCTATGCTTAAGAAAAAAGCAGATTGGSCCTTGGATGGATAGCAGATAGAAAA 540
Db 535 ACCATGCCCTATGCTTAAGAAAAAAGCAGATTGGSCCTTGGATGGATAGCAGATAGAAAA 594
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCCTTTTGGCTGTAGAAAGAGTTCCTTCTCAGCA 600
Db 595 TCTACTTTTGGGGAAGAGTGGTGGCCTTTTGGCTGTAGAAAGAGTTCCTTCTCAGCA 654
QY 601 TCTTTTGGCTGTATATTTCTGGCTAAAGAAAGAGAGGTCTTATGCCCAGACTCCTTTTTC 658
Db 655 TCTTTTGGCTGTATATTTCTGGCTAAAGAAAGAGAGGTCTTATGCCCAGACTCCTTTTTC 714
QY 659 CCAATGAACCTCATCAGCAGA 678
Db 715 CAATGAACCTCATCAGCAGA 734
```

Search completed: October 30, 2005, 03:23:17
Job time : 2721.5 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	1053	100.0	1053	6	AR454868	AR454868 Sequence
2	1053	100.0	1053	6	BD064764	BD064764 Novel pro
3	1053	100.0	1053	6	BD093077	BD093077 Novel pro
4	1053	100.0	1056	6	CQ714252	CQ714252 Sequence
5	1053	100.0	1056	9	AB166671	AB166671 Homo sapi
6	1053	100.0	4955	6	AR454869	AR454869 Sequence
7	1053	100.0	4955	6	BD064765	BD064765 Novel pro
8	1053	100.0	4955	6	BD093078	BD093078 Novel pro
9	1053	100.0	4955	9	AR036063	AR036063 Homo sapi
10	1051.4	99.8	1053	6	AR454877	AR454877 Sequence
11	1051.4	99.8	1053	6	BD064773	BD064773 Novel pro
12	1051.4	99.8	1053	6	BD093086	BD093086 Novel pro
13	1051.4	99.8	1081	6	AR454870	AR454870 Sequence
14	1051.4	99.8	1081	6	BD064766	BD064766 Novel pro
15	1051.4	98.8	1081	6	BD093079	BD093079 Novel pro
16	1049.8	99.7	1601	6	BD156916	BD156916 Primer fo
17	1049.8	99.7	1601	6	AX877905	AX877905 Sequence
18	1049.8	99.7	1601	9	AK001965	AK001965 Homo sapi
19	895.2	85.0	4532	10	BC058103	BC058103 Mus musci


```
Db 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT 300
Qy 301 GGAATTTGAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGGTTCCAGAGCT 360
Db 301 GGAATTTGAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGGTTCCAGAGCT 360
Qy 361 CGCTGTTTCTATGCTTTCAAATTTCTCATCAGAGATGTTCACTCAGAGATGTCAGATTG 420
Db 361 CGCTGTTTCTATGCTTTCAAATTTCTCATCAGAGATGTTCACTCAGAGATGTCAGATTG 420
Qy 421 CTGATAGACACTTATCATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA 480
Db 421 CTGATAGACACTTATCATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAAGCAGATGGGCGCTTGCATGATGATGACAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAGCAGATGGGCGCTTGCATGATGATGACAGATAGAAA 540
Qy 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTCTGCTGTAGAGAGTCTTATGCCAGGACTCAC 600
Db 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTCTGCTGTAGAGAGTCTTATGCCAGGACTCAC 600
Qy 601 TCTTTTGTGCTATATTTCTGCTTAAAGAGAGAGTCTTATGCCAGGACTCACATTTTTC 660
Db 601 TCTTTTGTGCTATATTTCTGCTTAAAGAGAGAGTCTTATGCCAGGACTCACATTTTTC 660
Qy 661 AATGAATCTATCAGAGATGAGAGTCTTCACTGTGATCTTGTGCTGATGTTCCAA 720
Db 661 AATGAATCTATCAGAGATGAGAGTCTTCACTGTGATCTTGTGCTGATGTTCCAA 720
Qy 721 TACTTAGTAAATAGCCTTCAAGAAAAGGTCAGGAGATCATTTGTGATGCTGTCAA 780
Db 721 TACTTAGTAAATAGCCTTCAAGAAAAGGTCAGGAGATCATTTGTGATGCTGTCAA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCCAGTGGCCCTCATTTGGAATGAAATTGCA 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCCAGTGGCCCTCATTTGGAATGAAATTGCA 840
Qy 841 TTGATGAAAAGTACATGAGTTTGTAGTCAGAGATTAATGTTGGAACTTGGATTTCTCA 900
Db 841 TTGATGAAAAGTACATGAGTTTGTAGTCAGAGATTAATGTTGGAACTTGGATTTCTCA 900
Qy 901 AAGGTTTTCAGGAGAAAATCCCTTTGATTTTATGGAACATTTCTTTAGAGGAAA 960
Db 901 AAGGTTTTCAGGAGAAAATCCCTTTGATTTTATGGAACATTTCTTTAGAGGAAA 960
Qy 961 ACAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTGTGAGTTATGGCAGAAAC 1020
Db 961 ACAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTGTGAGTTATGGCAGAAAC 1020
Qy 1021 ACAGATAACGTTTCACTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGTTTCACTTGGATGAGATTTT 1053
```

```
RESULT 2
BD064764
LOCUS BD064764
DEFINITION Novel protein and DNA thereof.
ACCESSION BD064764
VERSION BD064764.1 GI:22610367
KEYWORDS JP 2001269184-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1053)
Nakamura,Y., Arakawa,H. and Tanaka,H.
Novel protein and DNA thereof
Patent: JP 2001269184-A 1 02-Oct-2001;
YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD
OS Homo sapiens (human)
```

```
PN JP 2001269184-A/1
PD 02-OCT-2001
PF 27-JUN-2000 JP 2000192401
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N15/09, A61K31/711, A61K38/00, A61K39/43, A61K45/00, A61K48/00, PC
A61P35/00,
PC A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC
, C12N5/10, C12N9/02,
PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53// (C12N15/09,
PC C12R1/91), (C12N15/09, C12R1/19), (C12N9/02, C12R1/91), (C12P21/02, C12R1/91),
PC (C12P21/08, C12R1/91), C12N15/00, A61K37/02, A61K37/48, C12N5/00,
PC (C12N15/00, C12R1/91), (C12N15/00, C12R1/19)
CC Novel protein and DNA thereof
FH Key Location/Qualifiers
FT source 1..1053
FT /organism="Homo sapiens (human)".
FEATURES
source
1..1053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 1053; DB 6; Length 1053;
Best Local Similarity 100.0%; Pred. No. 1.2e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGGCGACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGGCGACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Qy 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGGCCACTCTCTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGGCCACTCTCTAAGAAAGAGTTCTCGC 120
Qy 121 CGGTTGTCTATCTTTCCAAATCCAGTACCTCATATTTGGAATAATGTATAACAGGCACAG 180
Db 121 CGGTTGTCTATCTTTCCAAATCCAGTACCTCATATTTGGAATAATGTATAACAGGCACAG 180
Qy 181 GCTTCTCTTCTGGAACAGCAGAGAGGTGAGTTATCAAGAGATCTCCCTCATCTGGAACAG 240
Db 181 GCTTCTCTTCTGGAACAGCAGAGAGGTGAGTTATCAAGAGATCTCCCTCATCTGGAACAG 240
Qy 241 CTTAAAGCAGATGAGAGTACTTCTATCTCTCATCTTAGCCTTTTTCAGCCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAGTACTTCTATCTCTCATCTTAGCCTTTTTCAGCCAGTGAT 300
Qy 301 GGAATTTGAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGGTTCCAGAGCT 360
Db 301 GGAATTTGAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGGTTCCAGAGCT 360
Qy 361 CGCTGTTTCTATGCTTTCAAATTTCTCATCAGAGATGTTTCACTCAGAGATGTACAGTTG 420
Db 361 CGCTGTTTCTATGCTTTCAAATTTCTCATCAGAGATGTTTCACTCAGAGATGTACAGTTG 420
Qy 421 CTGATAGACACTTATCATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA 480
Db 421 CTGATAGACACTTATCATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAAGCAGATGGGCGCTTGCATGATGATGACAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAGCAGATGGGCGCTTGCATGATGATGACAGATAGAAA 540
Qy 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTCTGCTGTAGAGAGTCTTATGCCAGGACTCAC 600
Db 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTCTGCTGTAGAGAGTCTTATGCCAGGACTCAC 600
Qy 601 TCTTTTGTGCTATATTTCTGCTTAAAGAGAGTCTTATGCCAGGACTCACATTTTTC 660
Db 601 TCTTTTGTGCTATATTTCTGCTTAAAGAGAGTCTTATGCCAGGACTCACATTTTTC 660
Qy 661 AATGAATCTATCAGAGATGAGAGTCTTCACTGTGATCTTGTGCTGATGTTCCAA 720
Db 661 AATGAATCTATCAGAGATGAGAGTCTTCACTGTGATCTTGTGCTGATGTTCCAA 720
Qy 721 TACTTAGTAAATAGCCTTCAAGAAAAGGTCAGGAGATCATTTGTGATGCTGTCAA 780
Db 721 TACTTAGTAAATAGCCTTCAAGAAAAGGTCAGGAGATCATTTGTGATGCTGTCAA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCCAGTGGCCCTCATTTGGAATGAAATTGCA 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCCAGTGGCCCTCATTTGGAATGAAATTGCA 840
Qy 841 TTGATGAAAAGTACATGAGTTTGTAGTCAGAGATTAATGTTGGAACTTGGATTTCTCA 900
Db 841 TTGATGAAAAGTACATGAGTTTGTAGTCAGAGATTAATGTTGGAACTTGGATTTCTCA 900
Qy 901 AAGGTTTTCAGGAGAAAATCCCTTTGATTTTATGGAACATTTCTTTAGAGGAAA 960
Db 901 AAGGTTTTCAGGAGAAAATCCCTTTGATTTTATGGAACATTTCTTTAGAGGAAA 960
Qy 961 ACAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTGTGAGTTATGGCAGAAAC 1020
Db 961 ACAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTGTGAGTTATGGCAGAAAC 1020
Qy 1021 ACAGATAACGTTTCACTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGTTTCACTTGGATGAGATTTT 1053
```



```

Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCTGCTGATGTTCCAA 720
Qy 721 TACTTAGTAAATAAGCCTTTCAGAGAAAGGCTCAGGAGAGATCATTTGTTGATGCTGTCAAA 780
Db 721 TACTTAGTAAATAAGCCTTTCAGAGAAAGGCTCAGGAGAGATCATTTGTTGATGCTGTCAAA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGCCAGTTCGCTCATTTGGGAATGAATTCGATT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGCCAGTTCGCTCATTTGGGAATGAATTCGATT 840
Qy 841 TTGATGAACAGTACATTTGATGTTGTAGCTACAGATTTACTTTGGAACTTTGGATTCTCA 900
Db 841 TTGATGAACAGTACATTTGATGTTGTAGCTACAGATTTACTTTGGAACTTTGGATTCTCA 900
Qy 901 AAGGTTTTTTCAGGACAGAAATCCTTTTCATTTTATGAAACATTTCTTTTGAAGGAAAA 960
Db 901 AAGGTTTTTTCAGGACAGAAATCCTTTTCATTTTATGAAACATTTCTTTTGAAGGAAAA 960
Qy 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGGTTTTGCAAGTTATGCGCAGAAAC 1020
Db 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGGTTTTGCAAGTTATGCGCAGAAAC 1020
Qy 1021 ACAGATAAGCTTTCACCTTGGATGCGAGTTT 1053
Db 1021 ACAGATAAGCTTTCACCTTGGATGCGAGTTT 1053

```

```

RESULT 3
BD093077
LOCUS Novel protein and its DNA. linear PAT 27-AUG-2002
DEFINITION BD093077
ACCESSION BD093077
VERSION BD093077.1 GI:22638665
KEYWORDS WO 0100799-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1053)
AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE Novel protein and its DNA
JOURNAL Patent: WO 0100799-A 1 04-JAN-2001;
TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,
HIROSHI TANAKA
COMMENT OS Homo sapiens (human)
PN WO 0100799-A/1
PD 04-JAN-2001
PF 27-JUN-2000 WO 2000JP004189
PR 28-JUN-1999 JP 99P 181131,06-JUL-1999 JP 99P 192391 PR
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N1/04, C12N1/53, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC
C12P21/02, A61K38/44,
PC A61K48/00, C07K16/40, G01N33/53, C12Q1/26
CC
FH Key Location/Qualifiers.

```

```

FEATURES
source
1. .1053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 1053; DB 6; Length 1053;
Beet Local Similarity 100.0%; Pred. No. 1.2e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGCGACCGGAAAGCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGCGACCGGAAAGCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Qy 61 GACACCAAGAAAGTGAATAAGTCAATGAAGAGCCACTCTCAAGAAAGAGTTCTGCG 120

```

```

FEATURES
source
1. .1053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 1053; DB 6; Length 1053;
Beet Local Similarity 100.0%; Pred. No. 1.2e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGCGACCGGAAAGCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGCGACCGGAAAGCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Qy 61 GACACCAAGAAAGTGAATAAGTCAATGAAGAGCCACTCTCAAGAAAGAGTTCTGCG 120

```

```

RESULT 4
CQ714252
LOCUS Sequence 186 from Patent WO02068579. linear PAT 03-FEB-2004
DEFINITION CQ714252
ACCESSION CQ714252

```

```

Db 61 GACACCAAGAAAGTGAATAAGTCAATGAAGAGCCACTCTCAAGAAAGAGTTCTGCG 120
Qy 121 CGGTTTGTCTCTTTCCAAATCCAGTACCTGATATTTGGAAAAATGATAAAACAGGCACAG 180
Db 121 CGGTTTGTCTCTTTCCAAATCCAGTACCTGATATTTGGAAAAATGATAAAACAGGCACAG 180
Qy 181 GCTTCTTCTGACAGCAGAGAGGTGCGATTTATCAAGAGATCTCCTCCTCAGTGGAAACAG 240
Db 181 GCTTCTTCTGACAGCAGAGAGGTGCGATTTATCAAGAGATCTCCTCCTCAGTGGAAACAG 240
Qy 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCAATCTTTAGCCTTTTTCAGGCGCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCAATCTTTAGCCTTTTTCAGGCGCAGTGAT 300
Qy 301 GGAATCTTAATGAATAATTTGGTGGAGCGCTTTAGTCAAGGAGTGCAGGTCCAGAGCT 360
Db 301 GGAATCTTAATGAATAATTTGGTGGAGCGCTTTAGTCAAGGAGTGCAGGTCCAGAGCT 360
Qy 361 CGCTGTTTCTATGCTTTCAAAATTTCTCATCAGAGATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGCTTTCAAAATTTCTCATCAGAGATGTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACACTTATCATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATGAA 480
Db 421 CTGATAGACACTTATCATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTGCATGATGAGCAGATGAAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTGCATGATGAGCAGATGAAAA 540
Qy 541 TCTACTTTTGGGAAAGAGTGTGCGCTTTCTGCTGTAGAGAGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGTGCGCTTTCTGCTGTAGAGAGAGTTTCTTCTCAGGA 600
Qy 601 TCTTTTGTCTGCTATATTTGCTTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Db 601 TCTTTTGTCTGCTATATTTGCTTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Qy 661 AATGAACCTCATCAGCAGAGATGAAGGCTTCACTGTGACTTTGCTGCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGGCTTCACTGTGACTTTGCTGCTGATGTTCCAA 720
Qy 721 TACTTAGTAAATAAGCCTTTCAGAGAAAGGCTCAGGAGAGATCATTTGTTGATGCTGCA 780
Db 721 TACTTAGTAAATAAGCCTTTCAGAGAAAGGCTCAGGAGAGATCATTTGTTGATGCTGCA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGCCAGTTCGCTCATTTGGAATGAATTCGATT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGCCAGTTCGCTCATTTGGAATGAATTCGATT 840
Qy 841 TTGATGAACAGTACATTTGATGTTGTAGCTACAGATTTACTTTGGAACTTTGGATTCTCA 900
Db 841 TTGATGAACAGTACATTTGATGTTGTAGCTACAGATTTACTTTGGAACTTTGGATTCTCA 900
Qy 901 AAGGTTTTTTCAGGACAGAAATCCTTTTCATTTTATGAAACATTTCTTTTGAAGGAAAA 960
Db 901 AAGGTTTTTTCAGGACAGAAATCCTTTTCATTTTATGAAACATTTCTTTTGAAGGAAAA 960
Qy 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGGTTTTGCAAGTTATGCGCAGAAAC 1020
Db 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGGTTTTGCAAGTTATGCGCAGAAAC 1020
Qy 1021 ACAGATAAGCTTTCACCTTGGATGCGAGTTT 1053
Db 1021 ACAGATAAGCTTTCACCTTGGATGCGAGTTT 1053

```

```

RESULT 4
CQ714252
LOCUS Sequence 186 from Patent WO02068579. linear PAT 03-FEB-2004
DEFINITION CQ714252
ACCESSION CQ714252

```

```

VERSION      CQ714252.1  GI:42275109
KEYWORDS
SOURCE       Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS      Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE        Kits, such as nucleic acid arrays, comprising a majority of
              humanexons or transcripts, for detecting expression and other uses
              thereof
JOURNAL      Patent: WO 02068579-A 186 06-SEP-2002;
              PE Corporation (NY) (US)
FEATURES
source       Location/Qualifiers
              1..1056
              /organism="Homo sapiens"
              /mol_type="unassigned DNA"
              /db_xref="taxon:9606"
ORIGIN
Query Match      100.0%; Score 1053; DB 6; Length 1056;
Best Local Similarity 100.0%; Pred. No. 1.2e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGGCGACCGGAAAGCGCGGCGGCTGGATCAGATGAGATGATCATCTTCA 60
Db 1 ATGGGCGACCGGAAAGCGCGGCGGCTGGATCAGATGAGATGATCATCTTCA 60
Qy 61 GACACCAACGAACTGAATTAAGTCAATTAAGAGCCACTCCTTAAGAAAGATTCTCGC 120
Db 61 GACACCAACGAACTGAATTAAGTCAATTAAGAGCCACTCCTTAAGAAAGATTCTCGC 120
Qy 121 CGGTTCGTCTATCTTCCAAATCCAGTACCTCATATTTGGAAATGTTAAACAGGCACAG 180
Db 121 CGGTTCGTCTATCTTCCAAATCCAGTACCTCATATTTGGAAATGTTAAACAGGCACAG 180
Qy 181 GCTTCCTCTCGACAGCAGAGAGGTGCAGTTATCAAGAGGATCTCCCTCACTGGAACAAG 240
Db 181 GCTTCCTCTCGACAGCAGAGAGGTGCAGTTATCAAGAGGATCTCCCTCACTGGAACAAG 240
Qy 241 CTTAAACAGATGAGAGTACTTCACTCTCACATCTTAGCTTTTTCGAGCCAGTGAT 300
Db 241 CTTAAACAGATGAGAGTACTTCACTCTCACATCTTAGCTTTTTCGAGCCAGTGAT 300
Qy 301 GGAATGTTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
Db 301 GGAATGTTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
Qy 361 CGCTGTTTCTATGCGCTTTCAAATTTCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGCGCTTTCAAATTTCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAPAGCAGATTGGCGCTTCGATGATGATGATAGAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAPAGCAGATTGGCGCTTCGATGATGATGATAGAGAAA 540
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
Qy 601 TCTTTTCTGCTATATTTCTGGCTAAAGAGAGGTTCTTATGCCAGGACTCACTTTTCC 660
Db 601 TCTTTTCTGCTATATTTCTGGCTAAAGAGAGGTTCTTATGCCAGGACTCACTTTTCC 660
Qy 661 AATGAATCATCAGCAGATGAGGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 720
Db 661 AATGAATCATCAGCAGATGAGGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 720
Qy 721 TACTTAGTAAATAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db 721 TACTTAGTAAATAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAAAGCCCTTGCAGTTGGCCCTCATTTGGAATGAATTGCATT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAAAGCCCTTGCAGTTGGCCCTCATTTGGAATGAATTGCATT 840
Qy 841 TTGATGAAAACAGTACATTTGAGTTTGTAGCTGACAGATTTACTTTGTGAACTTGGATTCTCA 900
Db 841 TTGATGAAAACAGTACATTTGAGTTTGTAGCTGACAGATTTACTTTGTGAACTTGGATTCTCA 900
Qy 901 AAGGTTTTTTCAGGCAGAAAAATCCTTTTGATTTTATGAAAAACATTTCTTTAGAGAAAA 960
Db 901 AAGGTTTTTTCAGGCAGAAAAATCCTTTTGATTTTATGAAAAACATTTCTTTAGAGAAAA 960
Qy 961 ACAAATTTCTTTTGAGAAACGAGTTTTCAGAGTATCAGCGCTTTTGCAGTTATGTCAGAAACC 1020
Db 961 ACAAATTTCTTTTGAGAAACGAGTTTTCAGAGTATCAGCGCTTTTGCAGTTATGTCAGAAACC 1020
Qy 1021 ACAGATAACGTCCTTTCACCTTCGATGCGAGATTTT 1053
Db 1021 ACAGATAACGTCCTTTCACCTTCGATGCGAGATTTT 1053
Qy 1056 bp mRNA linear PRI 09-MAR-2004
AB166671 Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase
small subunit 2, complete cds.
ACCESSION AB166671
VERSION AB166671.1 GI:45259568
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          1
          Ugai, H. and Yokoyama, K. K.
          Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase
          small subunit 2, complete cds
          Published Only in Database (2004)
          2 (bases 1 to 1056)
          Ugai, H. and Yokoyama, K. K.
          Direct Submission
          Submitted (05-MAR-2004) Hideyo Ugai, RIKEN Bioresource Center, Gene
          Engineering Division; 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074,
          Japan (E-mail:ugai@rtc.riken.go.jp, Tel:81-29-836-3612,
          Fax:81-29-836-9120)
          Location/Qualifiers
          1..1056
          /organism="Homo sapiens"
          /mol_type="mRNA"
          /db_xref="taxon:9606"
          /clone="p53R2.C19"
          /cell_line="Saos-2"
          /cell_type="Osteogenic sarcoma"
          1..1056
          /gene="p53R2"
          1..1056
          /gene="p53R2"
          /codon_start=1
          /product="p53-inducible ribonucleotide reductase small
          subunit 2"
          /protein_id="BAD12267.1"
          /db_xref="GI:45259568"
          /translation="MGDPERPEAAAGLDQDERSSDTNESEIKSNEEPLIKSSRRFVI
          PFIQYDPIWKYKQASFTABEVDLSKDLPHWNKLKADSKYFTSHLAPFAASDGI
          VNENLVERFQEVQVPEARCFYQIILNIENVHSEMYSLIDTYIDDPKRELFNAIE
          TMPYKKKADNRLNWIADKSTGFRVVFAPAANVAVGVFFSGFAALFWLXKGLMFLGT
          FSNELISRLGHLCDFACLMPQYLVNKPSEERVREIIVDAVKIIOEFLTEALPVLIG
          FNCLIMKQYIETFVADRLLELVELGFSKVAENPFPMENISLEGKTNFFEKRVSEYQF
          AVMAETDNDVFTLDADE"
FEATURES
source
          1..1056
          /organism="Homo sapiens"
          /mol_type="mRNA"
          /db_xref="taxon:9606"
          /clone="p53R2.C19"
          /cell_line="Saos-2"
          /cell_type="Osteogenic sarcoma"
          1..1056
          /gene="p53R2"
          1..1056
          /gene="p53R2"
          /codon_start=1
          /product="p53-inducible ribonucleotide reductase small
          subunit 2"
          /protein_id="BAD12267.1"
          /db_xref="GI:45259568"
          /translation="MGDPERPEAAAGLDQDERSSDTNESEIKSNEEPLIKSSRRFVI
          PFIQYDPIWKYKQASFTABEVDLSKDLPHWNKLKADSKYFTSHLAPFAASDGI
          VNENLVERFQEVQVPEARCFYQIILNIENVHSEMYSLIDTYIDDPKRELFNAIE
          TMPYKKKADNRLNWIADKSTGFRVVFAPAANVAVGVFFSGFAALFWLXKGLMFLGT
          FSNELISRLGHLCDFACLMPQYLVNKPSEERVREIIVDAVKIIOEFLTEALPVLIG
          FNCLIMKQYIETFVADRLLELVELGFSKVAENPFPMENISLEGKTNFFEKRVSEYQF
          AVMAETDNDVFTLDADE"
ORIGIN
```

Query Match 100.0%; Score 1053; DB 9; Length 1056;
Best Local Similarity 100.0%; Pred. No. 1.2e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCGACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
DB 1 ATGGGCGACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAAGAAAGATTCTCGC 120
DB 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAAGAAAGATTCTCGC 120
QY 121 CGGTTTGTCAATTTTCAATCCAGTACCTTGATATTTGGAAATGTATAACAGGCGCAG 180
DB 121 CGGTTTGTCAATTTTCAATCCAGTACCTTGATATTTGGAAATGTATAACAGGCGCAG 180
QY 181 GCTTCCTTCTGGACAGAGAGGTCGACTTATCAAGAGTCTCCCTCACTGGAACAG 240
DB 181 GCTTCCTTCTGGACAGAGAGGTCGACTTATCAAGAGTCTCCCTCACTGGAACAG 240
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCACATCTTAGCTTTTGGACGCGAT 300
DB 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCACATCTTAGCTTTTGGACGCGAT 300
QY 301 GGAATTGTAATGAATAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAAGTTTCCAGAGGCT 360
DB 301 GGAATTGTAATGAATAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAAGTTTCCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 420
DB 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 480
DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 480
QY 481 ACCATGCGCTATGTTAAGAAAAAGCAGATTTGGCCCTTGGATGATAGCAGATAGAAAA 540
DB 481 ACCATGCGCTATGTTAAGAAAAAGCAGATTTGGCCCTTGGATGATAGCAGATAGAAAA 540
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGTTTCTTCTCAGGA 600
DB 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGTTTCTTCTCAGGA 600
QY 601 TCTTTTGTCTGATATTTCTGGCTAAAGAGAGGTCCTTATGCCAGGACTCACTTTTTC 660
DB 601 TCTTTTGTCTGATATTTCTGGCTAAAGAGAGGTCCTTATGCCAGGACTCACTTTTTC 660
QY 661 AATGAACTCATCAGCAGATGAAGACTTTCATCTGTGACTTTGCTTGGCTGATGTTCCAA 720
DB 661 AATGAACTCATCAGCAGATGAAGACTTTCATCTGTGACTTTGCTTGGCTGATGTTCCAA 720
QY 721 TACTTAGTAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTGATGCTGTCAAA 780
DB 721 TACTTAGTAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTGATGCTGTCAAA 780
QY 781 ATTGAGCAGAGTTTAAACAGAGCCTTCCAGTTGGCCTCATTTGGAATGAATTCAT 840
DB 781 ATTGAGCAGAGTTTAAACAGAGCCTTCCAGTTGGCCTCATTTGGAATGAATTCAT 840
QY 841 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGGAACTTGGATTTCTCA 900
DB 841 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGGAACTTGGATTTCTCA 900
QY 901 AAGTTTTTTCAGGCAGAAAAATCCTTTTGAATTTTATGAAAAATTTTCTTTAGAGGAAAA 960
DB 901 AAGTTTTTTCAGGCAGAAAAATCCTTTTGAATTTTATGAAAAATTTTCTTTAGAGGAAAA 960
QY 961 ACAAATTTCTTGGAAACAGTTTTCAGAGTATCAGCGTTTTCAGTTTATGGCAGAAACC 1020
DB 961 ACAAATTTCTTGGAAACAGTTTTCAGAGTATCAGCGTTTTCAGTTTATGGCAGAAACC 1020

QY 1021 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1053
DB 1021 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1053

RESULT 6
AR454869
LOCUS AR454869
DEFINITION Sequence 3 from patent US 6682917.
ACCESSION AR454869
VERSION AR454869.1 GI:42688824
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4955)
AUTHORS Nakamura, Y., Arakawa, H. and Tanaka, H.
TITLE Protein having a ribonucleotide Reductase activity and a DNA thereof
JOURNAL Patent: US 6682917-A 3 27-JAN-2004;
FEATURES Location/Qualifiers
source 1. 4955
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 1053; DB 6; Length 4955;
Best Local Similarity 100.0%; Pred. No. 1.1e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY -1 ATGGGCGACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 245 ATGGGCGACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAAGAAAGATTCTCGC 120
DB 305 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAAGAAAGATTCTCGC 364
QY 121 CGGTTTGTCAATTTTCCAAATCCAGTACCTTGATATTTGGAAATGTATAACAGGCGCAG 180
DB 365 CGGTTTGTCAATTTTCCAAATCCAGTACCTTGATATTTGGAAATGTATAACAGGCGCAG 424
QY 181 GCTTCCTTCTGGACAGCAGAGAGTGCATTTATCAAAAGGATCTCCCTCACTGGAACAG 240
DB 425 GCTTCCTTCTGGACAGCAGAGAGTGCATTTATCAAAAGGATCTCCCTCACTGGAACAG 484
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTAGCGCTTTTTCGACGCGATGAT 300
DB 485 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTAGCGCTTTTTCGACGCGATGAT 544
QY 301 GGAATTGTAATGAATAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAAGTTTCCAGAGGCT 360
DB 545 GGAATTGTAATGAATAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAAGTTTCCAGAGGCT 604
QY 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTTG 420
DB 605 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTTGAA 480
DB 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTTGAA 724
QY 481 ACCATGCGCTATGTTAAGAAAAAGCAGATTGGGCTTTGCCATGATGATAGCAGATAGAAAA 540
DB 725 ACCATGCGCTATGTTAAGAAAAAGCAGATTGGGCTTTGCCATGATGATAGCAGATAGAAAA 784
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTCTTCTCTCAGGA 600
DB 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTCTTCTCTCAGGA 844
QY 601 TCTTTTGTCTGCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
DB 845 TCTTTTGTCTGCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 904


```
RESULT 8
BD093078
LOCUS      BD093078          4955 bp      DNA      linear      PAT 27-AUG-2002
DEFINITION Novel protein and its DNA.
ACCESSION  BD093078
VERSION     BD093078.1 GI:22638666
KEYWORDS    WO 0100799-A/2.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 4955)
AUTHORS     Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE       Novel protein and its DNA
JOURNAL     Patent: WO 0100799-A 2 04-JAN-2001;
            TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,
            HIROSHI TANAKA
COMMENT     OS Homo sapiens (human)
            PN WO 0100799-A/2
            PD 04-JAN-2001
            PE 27-JUN-2000 WO 2000JP004189
            PR 28-JUN-1999 JP 99P 181131,06-JUL-1999 JP 99P 192391 PR
            21-JAN-2000 JP 00P 017770
            PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
            C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N5/10 PC
            C12P21/02, A61K38/44,
            PC A61K48/00, C07K16/40, G01N33/53, C12Q1/26
            CC
FEATURES    source
            FH Key Location/Qualifiers.
            1..4955
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 1053; DB 6; Length 4955;
Best Local Similarity 100.0%; Pred. No. 1.1e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGGACCGGAAAGCGCGGAGCGCGGCTGGATCAGAGTGAAGATCATCTTCA 60
DB 1 ATGGCGGACCGGAAAGCGCGGAGCGCGGCTGGATCAGAGTGAAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGC 120
DB 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGC 364
QY 121 CGGTTTGTCTATCTTCCAAATCCAGTACCTGATATTGGAAATGTTAAACAGGCACAG 180
DB 365 CGGTTTGTCTATCTTCCAAATCCAGTACCTGATATTGGAAATGTTAAACAGGCACAG 424
QY 181 GCTTCTCTTGGCAGCAGAGAGTGCATCTTCAAAAGATCTCCCTCACTGGAACAAG 240
DB 425 GCTTCTCTTGGCAGCAGAGAGTGCATCTTCAAAAGATCTCCCTCACTGGAACAAG 484
QY 241 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCTTTTTCGAGCCAGTGAT 300
DB 485 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCTTTTTCGAGCCAGTGAT 544
QY 301 GGAATTTGAATGAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
DB 545 GGAATTTGAATGAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 604
QY 361 CGCTGTTTCTATGCTTTCAAAATCTCATCTGAGAAAGTTCACATCAGAGATGACATTTTC 420
DB 605 CGCTGTTTCTATGCTTTCAAAATCTCATCTGAGAAAGTTCACATCAGAGATGACATTTTC 664
QY 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTATGCAATTTGAA 480
DB 665 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTATGCAATTTGAA 724
```

```
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTTCGGATGGATAGCAGATAGAAA 540
DB 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTTCGGATGGATAGCAGATAGAAA 784
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTCGCTGCTGTAGAGAGGATTTTCTTCTCAGA 600
DB 785 TCTACTTTTGGGGAAGAGTGGTGGCTTTCGCTGCTGTAGAGAGGATTTTCTTCTCAGA 844
QY 601 TCTTTTGTCTATATCTTCTGCTAAGAGAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
DB 845 TCTTTTGTCTATATCTTCTGCTAAGAGAGAGAGTCTTATGCCAGGACTCACTTTTCC 904
QY 661 AATGAATCTCATCAGCAGAGATGAAGGACTTCACTCTGACTTTTGTCTGCTGATGTTCCAA 720
DB 905 AATGAATCTCATCAGCAGAGATGAAGGACTTCACTCTGACTTTTGTCTGCTGATGTTCCAA 964
QY 721 TACTTAGTAATAAGCCTTTCAGAAAGAGGCTCAGGAGATCATTTGATGATGCTGCTCAA 780
DB 965 TACTTAGTAATAAGCCTTTCAGAAAGAGGCTCAGGAGATCATTTGATGATGCTGCTCAA 1024
QY 781 ATTGAGCAGAGTTTAAACAGAGCCCTTCCAGTTGGCCTCATTTGGAATGAATTCGATT 840
DB 1025 ATTGAGCAGAGTTTAAACAGAGCCCTTCCAGTTGGCCTCATTTGGAATGAATTCGATT 1084
QY 841 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTTGTGGAATTTCTTCA 900
DB 1085 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTTGTGGAATTTCTTCA 1144
QY 901 AAGTTTTTTCAGGCAGAAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAAGAAA 960
DB 1145 AAGTTTTTTCAGGCAGAAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAAGAAA 1204
QY 961 ACAATTTTCTTTGAGAAACAGATTTCAGAGTATCAGCGTTTTCAGTTATGCGCAGAAACC 1020
DB 1205 ACAATTTTCTTTGAGAAACAGATTTCAGAGTATCAGCGTTTTCAGTTATGCGCAGAAACC 1264
QY 1021 ACAGATAAGCTTCTCACCCTTGGATGCGAGATTTT 1053
DB 1265 ACAGATAAGCTTCTCACCCTTGGATGCGAGATTTT 1297
```

```
RESULT 9
AB036063
LOCUS      AB036063          4955 bp      mRNA      linear      PRI 11-MAR-2000
DEFINITION Homo sapiens p53R2 mRNA for ribonucleotide reductase, complete cds.
ACCESSION  AB036063
VERSION     AB036063.1 GI:7229085
KEYWORDS    p53R2; ribonucleotide reductase.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (sites)
AUTHORS     Tanaka,H., Arakawa,H., Yamaguchi,T., Shiraishi,K., Fukuda,S.,
            Matsui,K., Takei,Y. and Nakamura,Y.
TITLE       A ribonucleotide reductase gene involved in a p53-dependent
            cell-cycle checkpoint for DNA damage
JOURNAL     Nature 404 (6773), 42-49 (2000)
MEDLINE     20179179
PUBMED      10716435
REFERENCE   2 (bases 1 to 4955)
AUTHORS     Tanaka,H., Arakawa,H. and Nakamura,Y.
TITLE       Direct Submission
JOURNAL     Submitted (18-DEC-1999) Yusuke Nakamura, University of Tokyo,
            Institute of Medical Science, Human Genome Center, Laboratory of
            Molecular Medicine; 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639,
            Japan (E-mail:yusuke@ims.u-tokyo.ac.jp, Tel:+81-3-5449-5372,
            Fax:+81-3-5449-5433)
FEATURES    source
            1..4955
            /organism="Homo sapiens"
            /mol_type="mRNA"
```

```
/db_xref="taxon:9606"
/chromosome="8"
/map="8q23.1"
1..4955
/gene="p53R2"
245..1300
/gene="p53R2"
/codon_start=1
/product="ribonucleotide reductase"
/protein_id="BAA92434.1"
/db_xref="GI:7229086"
/translation="MGDPPEPAAGLDQDERSSDTNESIKNEEPLLRKSRFFV
FPIDYIMKRYKAQASFWTAEEVDLSKLPWNKLKADKPYFISHLAFPAASDGI
TNPVVERFQVEQVPEARCFYQIILINVENSMYSLLDITRDPKRBFPLFNAIE
TMPVKKKADWALRWIADRKSTFCGERVVAFAAEGVPFSGSPAALFWLKRGLMPGLT
FSNELISRDGLCHDFACILMFOVLNKPSEBVEREILVDVKLEQEPLTEALPVLGIG
MNCILMKQYIEFVADRLILVELGSKVFQAEINFPDFMENISLEKTNFFEKRVSYQRP
AVMAETTDNVFTLDADF"
4933
/polyA_site
/gene="p53R2"
/feature="22 a nucleotides"
```

ORIGIN

```
Query Match 100.0%; Score 1053; DB 9; Length 4955;
Best Local Similarity 100.0%; Pred. No. 1.le-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGACCGGAAAGCGCGGAGCGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
Db |
Qy 245 ATGGCGACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 304
Db |
Qy 61 GACACCAAGCAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGAGTCTCGC 120
Db |
Qy 305 GACACCAAGCAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGAGTCTCGC 364
Db |
Qy 121 CGGTTTGTCATCTTCAATCCAGTACCTGATATTTGGAATATGTAACAGGCAAG 180
Db |
Qy 365 CGGTTTGTCATCTTCAATCCAGTACCTGATATTTGGAATATGTAACAGGCAAG 424
Db |
Qy 181 GCTTCCTCTGGACAGCAGAGAGTGCATTAACAGATCTCCCTCACTGGAACAG 240
Db |
Qy 425 GCTTCCTCTGGACAGCAGAGAGTGCATTAACAGATCTCCCTCACTGGAACAG 484
Db |
Qy 241 CTTAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCTTTTGGACGAGTAT 300
Db |
Qy 485 CTTAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCTTTTGGACGAGTAT 544
Db |
Qy 301 GGAATTGTAAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGCT 360
Db |
Qy 545 GGAATTGTAAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGCT 604
Db |
Qy 361 CGCTGTTTCTATGGCTTTCAATTTCTCATCAGAAATGTTTCACTCAGAGATGTAAGTTG 420
Db |
Qy 605 CGCTGTTTCTATGGCTTTCAATTTCTCATCAGAAATGTTTCACTCAGAGATGTAAGTTG 664
Db |
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 480
Db |
Qy 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 724
Db |
Qy 481 ACCATGCGCTATGTTAAGAAAAGAGATGGGCGCTTGGATGGATGAGCAGATAGAAA 540
Db |
Qy 725 ACCATGCGCTATGTTAAGAAAAGAGATGGGCGCTTGGATGGATGAGCAGATAGAAA 784
Db |
Qy 541 TCTACTTTTGGGAAAGAGTGGCTTTGCTGCTGTAGAGAGTGTTCCTCTCAGGA 600
Db |
Qy 785 TCTACTTTTGGGAAAGAGTGGCTTTGCTGCTGTAGAGAGTGTTCCTCTCAGGA 844
Db |
Qy 601 TCTTTTGTGCTATATTTGCTTAAGAGAGAGTCTTATGCCAGACTCACTTTTTC 660
Db |
Qy 845 TCTTTTGTGCTATATTTGCTTAAGAGAGAGTCTTATGCCAGACTCACTTTTTC 904
Db |
Qy 661 AATGAATCTCATCAGCAGATGAAGGACTTCACCTGTGACTTTTGTGCTGATGTTCCAA 720
Db |
```

```
Db 905 AATGAATCTCATCAGCAGAGATGAAGGACTTCACCTGTGACTTTTGTGCTGATGTTCCAA 964
Qy 721 TACTTAGTAAATAAGCCCTTCAGAAAGAGGCTCAGGAGATCATTTGTGATGCTGTCAAA 780
Db |
Qy 965 TACTTAGTAAATAAGCCCTTCAGAAAGAGGCTCAGGAGATCATTTGTGATGCTGTCAAA 1024
Qy 781 ATTGAGCAGGAGTGTAAACAGAAAGCCCTTGCCAGTTGGCCCTCATTTGGAATGAATGCAAT 840
Db |
Qy 1025 ATTGAGCAGGAGTGTAAACAGAAAGCCCTTGCCAGTTGGCCCTCATTTGGAATGAATGCAAT 1084
Qy 841 TTGATGAAAACAGTACATTTGAGTTTGTAGCTGACAGATTTAATTGTGGAATCTTGGAATCTCA 900
Db |
Qy 1085 TTGATGAAAACAGTACATTTGAGTTTGTAGCTGACAGATTTAATTGTGGAATCTTGGAATCTCA 1144
Qy 901 AAGCTTTTTCAGGCGAGAAAATCCCTTTGATTTTATGGAACACATTTCTTTAGAAGAAA 960
Db |
Qy 1145 AAGCTTTTTCAGGCGAGAAAATCCCTTTGATTTTATGGAACACATTTCTTTAGAAGAAA 1204
Qy 961 ACAAATTTCTTTGAGAAAACGAGTTTTCAGAGTATCAGCGTTTGTGAGTTATGTCAGAAACC 1020
Db |
Qy 1205 ACAAATTTCTTTGAGAAAACGAGTTTTCAGAGTATCAGCGTTTGTGAGTTATGTCAGAAACC 1264
Qy 1021 ACAGATAACGTCCTTCACTTGGATGCGAGATTTT 1053
Db 1265 ACAGATAACGTCCTTCACTTGGATGCGAGATTTT 1297

RESULT 10
AR454877
LOCUS AR454877 1053 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 12 from patent US 6682917.
ACCESSION AR454877
VERSION AR454877.1 GI:42688832
KEYWORDS
SOURCE Unknown.
ORGANISM Unkown.
REFERENCE 1 (bases 1 to 1053)
AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE Protein having a ribonucleotide Reductase activity and a DNA
JOURNAL Patent: US 6682917-A 12 27-JAN-2004;
FEATURES Location/Qualifiers
source 1..1053
/mol_type="genomic DNA"

ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 1053;
Best Local Similarity 99.9%; Pred. No. 3e-255;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGCGACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
Db |
Qy 61 GACACCAAGCAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGAGTCTCGC 120
Db |
Qy 61 GACACCAAGCAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGAGTCTCGC 120
Qy 121 CGGTTTGTCATCTTCAATCCAGTACCTGATATTTGGAATATGTAACAGGCAAG 180
Db |
Qy 121 CGGTTTGTCATCTTCAATCCAGTACCTGATATTTGGAATATGTAACAGGCAAG 180
Qy 181 GCTTCCTCTTGGACAGCAGAGAGTGCATTTATCAAGGATCTCCCTCACTGGAACAG 240
Db |
Qy 181 GCTTCCTCTTGGACAGCAGAGAGTGCATTTATCAAGGATCTCCCTCACTGGAACAG 240
Qy 241 CTTAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCTTTTGGACGAGTAT 300
Db |
Qy 241 CTTAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCTTTTGGACGAGTAT 300
Qy 301 GGAATTGTAAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGCT 360
```


Db	301	GGATGTTAAATGAAATTTGGTGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT	360
Qy	361	CGCTGTTTCTATGGCTTTCAAAATTCCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG	420
Db	361	CGCTGTTTCTATGGCTTTCAAAATTCCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG	420
Qy	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTAAATGCAATGAA	480
Db	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTAAATGCAATGAA	480
Qy	481	ACCATGCCCTATGTTAAGAAAAGCAGATTCGGGCTTCGATGATAGCAGATAGAAA	540
Db	481	ACCATGCCCTATGTTAAGAAAAGCAGATTCGGGCTTCGATGATAGCAGATAGAAA	540
Qy	541	TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGAGTTTCTTCTCAGGA	600
Db	541	TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGAGTTTCTTCTCAGGA	600
Qy	601	TCTTTGCTGCTATATTCTGGCTAAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC	660
Db	601	TCTTTGCTGCTATATTCTGGCTAAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC	660
Qy	661	AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTTCCAA	720
Db	661	AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTTCCAA	720
Qy	721	TACTTAGTAAATAGCCCTTCAGAGAAAGGCTCAGGAGATCATTTGATGCTGTCAAA	780
Db	721	TACTTAGTAAATAGCCCTTCAGAGAAAGGCTCAGGAGATCATTTGATGCTGTCAAA	780
Qy	781	ATTGAGCAGGAGTTTAAACAGAGCCCTTGCCAGTTGGCTCATTTGGAATGCAATT	840
Db	781	ATTGAGCAGGAGTTTAAACAGAGCCCTTGCCAGTTGGCTCATTTGGAATGCAATT	840
Qy	841	TGATGAAACAGTACATTTGAGTTGTAGCTGACAGATTAATTGTTGGAATTTCTCA	900
Db	841	TGATGAAACAGTACATTTGAGTTGTAGCTGACAGATTAATTGTTGGAATTTCTCA	900
Qy	901	AAGTTTTTTCAGGAGAAATCCCTTTGATTTTATGGAACATTTCTTTAGAGGAAA	960
Db	901	AAGTTTTTTCAGGAGAAATCCCTTTGATTTTATGGAACATTTCTTTAGAGGAAA	960
Qy	961	ACAAATTTCTTTGAGAACAGATTTTACAGATATCAGCGTTTTCAGTTATGCGAGAAC	1020
Db	961	ACAAATTTCTTTGAGAACAGATTTTACAGATATCAGCGTTTTCAGTTATGCGAGAAC	1020
Qy	1021	ACAGATAACGCTTTCACTTGGATGCGATTTT	1053
Db	1021	ACAGATAACGCTTTCACTTGGATGCGATTTT	1053
RESULT 11			
BD064773			
LOCUS			
DEFINITION Novel protein and DNA thereof.			
ACCESSION BD064773			
VERSION BD064773.1 GI:22610376			
KEYWORDS JP 2001269184-A/10.			
SOURCE Homo sapiens (human)			
ORGANISM Homo sapiens			
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
1 (bases 1 to 1053)			
AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.			
TITLE Novel protein and DNA thereof			
JOURNAL Patent: JP 2001269184-A 10 02-OCT-2001;			
YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD			
OS Homo sapiens (human)			
PN JP 2001269184-A/10			
PD 02-OCT-2001			
PF 27-JUN-2000 JP 2000192401			
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC			

C12N15/09, A61K31/711, A61K38/00, A61K38/43, A61K45/00, A61K48/00, PC			
A61P35/00,			
PC A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC			
C12N5/10, C12N9/02,			
PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53// (C12N15/09,			
PC C12R1:91),			
PC (C12N15/09, C12R1:19), (C12N9/02, C12R1:91), (C12P21/02, C12R1:91),			
PC (C12P21/08, C12R1:91), (C12N15/00, A61K37/02, A61K37/48, C12N5/00,			
PC (C12N15/00, C12R1:91), (C12N15/00, C12R1:19)			
CC Novel protein and DNA thereof			
FH Key Location/Qualifiers			
FT Source 1..1053 /organism='Homo sapiens (human)'			
FEATURES			
Source 1..1053			
/organism='Homo sapiens'			
/mol_type='genomic DNA'			
/db_xref='taxon:9606'			
ORIGIN			
Query Match 99.8%; Score 1051.4; DB 6; Length 1053;			
Best Local Similarity 99.9%; Pred. No. 3e-255; 1; Indels 0; Gaps 0;			
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	1	ATGGCGGACCCGGAAGCCGGAAGCGCGCGCTGGATCAGGATCAGAGATCATCTTCA	60
Db	1	ATGGCGGACCCGGAAGCCGGAAGCGCGCGCTGGATCAGGATCAGAGATCATCTTCA	60
Qy	61	GACACCAACGAAGTGAATAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC	120
Db	61	GACACCAACGAAGTGAATAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC	120
Qy	121	CGGTTTGTCTTCCATCCAGTACCTGTATTTGGAAATGTATAAAGGACAG	180
Db	121	CGGTTTGTCTTCCATCCAGTACCTGTATTTGGAAATGTATAAAGGACAG	180
Qy	181	GCTTCTTCTGACAGCAGAGGTGCACTTATCAAGAGATCTCCCTCACTCGAACAAG	240
Db	181	GCTTCTTCTGACAGCAGAGGTGCACTTATCAAGAGATCTCCCTCACTCGAACAAG	240
Qy	241	CTTAAGCAGATGAGAAGTACTTCTCATCTCTCACTTTAGCCCTTTTTCGAGCCAGTGAT	300
Db	241	CTTAAGCAGATGAGAAGTACTTCTCATCTCTCACTTTAGCCCTTTTTCGAGCCAGTGAT	300
Qy	301	GGAATTTAAATGAAAATTTGGTGAGCGCTTTAGTGCAGAGGTGCGAGTTCCAGAGCT	360
Db	301	GGAATTTAAATGAAAATTTGGTGAGCGCTTTAGTGCAGAGGTGCGAGTTCCAGAGCT	360
Qy	361	CGCTGTTTCTATGGCTTTCAAAATTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG	420
Db	361	CGCTGTTTCTATGGCTTTCAAAATTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG	420
Qy	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA	480
Db	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA	480
Qy	481	ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTCGCGATGAGATAGAGATAAAA	540
Db	481	ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTCGCGATGAGATAGAGATAAAA	540
Qy	541	TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGAGTTTCTTCTCAGGA	600
Db	541	TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGAGTTTCTTCTCAGGA	600
Qy	601	TCTTTTGTGCTATATTCTGGCTAAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC	660
Db	601	TCTTTTGTGCTATATTCTGGCTAAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC	660
Qy	661	AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTTCCAA	720
Db	661	AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTTCCAA	720
Qy	721	TACTTAGTAAATAGCCCTTCAGAGAAAGGCTCAGGAGATCATTTGATGCTGTCAAA	780


```
Db 721 TACTTAGTAATAAGCCTTTCAGAGAAAGGCTCAGGAGATCAATGTTGTCGTCTCAAA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGCCAGTTGSCCTCAATGGAAATGAATTCATT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGCCAGTTGSCCTCAATGGAAATGAATTCATT 840
Qy 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACTTTGGATTCTCA 900
Db 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACTTTGGATTCTCA 900
Qy 901 AAGTTTTTTCAGGAGAAATCCCTTTGATTTTATGAAACATTTCTTTAGAGGAAAA 960
Db 901 AAGTTTTTTCAGGAGAAATCCCTTTGATTTTATGAAACATTTCTTTAGAGGAAAA 960
Qy 961 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1020
Db 961 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1020
Qy 1021 ACAGATAACGTCCTTCCACCTTCGATGCAGATTTT 1053
Db 1021 ACAGATAACGTCCTTCCACCTTCGATGCAGATTTT 1053

RESULT 12
LOCUS BD093086 1053 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel protein and its DNA.
ACCESSION BD093086
VERSION BD093086.1 GI:22638674
KEYWORDS WO 0100799-A/10.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1053)
AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE Novel protein and its DNA
JOURNAL Patent: WO 0100799-A 10 04-JAN-2001;
TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,
HIROSHI TANAKA
COMMENT OS Homo sapiens (human)
PN WO 0100799-A/10
PF 27-JUN-2000 WO 2000JP004189
PR 28-JUN-1999 JP 99P 181131,06-JUL-1999 JP 99P 192391 PR
21-JAN-2000 JP 00P 017770
PI YUSUKE NAKAMURA,HIROFUMI ARAKAWA,HIROSHI TANAKA PC
C12N9/04,C12N15/53,C12N1/15,C12N1/19,C12N1/21,C12N5/10 PC
,C12P21/02,A61K38/44,
PC A61K48/00,C07K16/40,G01N33/53,C12Q1/26
CC

FEATURES
source
FH Key Location/Qualifiers.
1..1053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 1053;
Best Local Similarity 99.9%; Pred. No. 3e-255;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGGCGACCGGAAAGCCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGGCGACCGGAAAGCCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Qy 61 GACACCAACGAACTGAAATAAATCAATGAAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAACTGAAATAAATCAATGAAGCCACTCTTAAGAAAGAGTTCTCGC 120
Qy 121 CGGTTTGTCTTCCAAATCCAGTACCCTCGATATTTGGAAATGTATAAACAGGCACAG 180
```

```
Db 121 CGGTTTGTCTTCCAAATCCAGTACCCTCGATATTTGGAAATGTATAAACAGGCACAG 180
Qy 181 GCTTCTCTCTGAGCAGCAGAGAGGTGCATTTATCABAAGGATCTCCCTCACTGGACACAG 240
Db 181 GCTTCTCTCTGAGCAGCAGAGAGGTGCATTTATCABAAGGATCTCCCTCACTGGACACAG 240
Qy 241 CTTAAAGCAGATGAGAGTACTTTCATCTCTCACTCTTAGCCTTTTTCAGCAGCAGTGTAT 300
Db 241 CTTAAAGCAGATGAGAGTACTTTCATCTCTCACTCTTAGCCTTTTTCAGCAGCAGTGTAT 300
Qy 301 GGAATGTGAAATGAAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
Db 301 GGAATGTGAAATGAAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
Qy 361 CGCTGTTTCTATGCTTCCAAATTTCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGCTTCCAAATTTCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATCAATTTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATCAATTTGAA 480
Qy 481 ACCATGCGCTATGTTAAGAAAAAAGCAGATTGGCGCTTTCGATGGATAGCAGATAGAAA 540
Db 481 ACCATGCGCTATGTTAAGAAAAAAGCAGATTGGCGCTTTCGATGGATAGCAGATAGAAA 540
Qy 541 TCTACTTTTGGGAAAGAGTGCTGGCTTTCGCTGCTGTAGAAGGATTTTCTTCTCAGA 600
Db 541 TCTACTTTTGGGAAAGAGTGCTGGCTTTCGCTGCTGTAGAAGGATTTTCTTCTCAGA 600
Qy 601 TCTTTTGTGCTATATTCTGGCTAAAGAAAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
Db 601 TCTTTTGTGCTATATTCTGGCTAAAGAAAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
Qy 661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGTGATGTTCCAA 720
Db 661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGTGATGTTCCAA 720
Qy 721 TACTTAGTAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAAA 780
Db 721 TACTTAGTAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAAA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGCCAGTTGCGCTCATTTGGAAATGAATTCATT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGCCAGTTGCGCTCATTTGGAAATGAATTCATT 840
Qy 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACTTTGGATTCTCA 900
Db 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACTTTGGATTCTCA 900
Qy 901 AAGTTTTTTCAGGAGAAATCCCTTTGATTTTATGAAACATTTCTTTAGAGGAAAA 960
Db 901 AAGTTTTTTCAGGAGAAATCCCTTTGATTTTATGAAACATTTCTTTAGAGGAAAA 960
Qy 961 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1020
Db 961 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1020
Qy 1021 ACAGATAACGTCCTTCCACCTTCGATGCAGATTTT 1053
Db 1021 ACAGATAACGTCCTTCCACCTTCGATGCAGATTTT 1053

RESULT 13
LOCUS AR454870 1081 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 4 from patent US 6682917.
ACCESSION AR454870
VERSION AR454870.1 GI:42688825
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
```

```
Unclassified.
REFERENCE 1 (bases 1 to 1081)
AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE Protein having a ribonucleotide Reductase activity and a DNA
        thereof
JOURNAL Patent: US 6682917-A 4 27-JAN-2004;
FEATURES   Location/Qualifiers
            source
            1..1081
            /organism="unknown"
            /mol_type="genomic DNA"

ORIGIN
Query Match      99.8%; Score 1051.4; DB 6; Length 1081;
Best Local Similarity 99.9%; Pred. No. 3e-255;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCGGAAAGCGCGGAGCGGCTGGATCAGGATGAGATGAGATCATCTTCA 60
DB 1 |
QY 20 ATGGGCGACCGGAAAGCGCGGAGCGGCTGGATCAGGATGAGATGAGATCATCTTCA 79
DB 2 |
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGC 120
DB 6 |
QY 80 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGC 139
DB 8 |
QY 121 CGGTTTGTCATCTTTCCATCCAGTACCCCTGATATTTGGAAATGTATAACAGGACAG 180
DB 1 |
QY 140 CGGTTTGTCATCTTTCCATCCAGTACCCCTGATATTTGGAAATGTATAACAGGACAG 199
DB 1 |
QY 181 GCTTCCTCTGGACAGAGAGGTCGATTAATCAAGAGTCTCCCTCACTGGAACAG 240
DB 1 |
QY 200 GCTTCCTCTGGACAGAGAGGTTGACTTATCAAGAGTCTCCCTCACTGGAACAG 259
DB 1 |
QY 241 CTTAAAGCAGATGAGAGTACTTCTCTCTCACATCTTAGCTTTTGGACGAGGAT 300
DB 1 |
QY 260 CTTAAAGCAGATGAGAGTACTTCTCTCTCACATCTTAGCTTTTGGACGAGGAT 319
DB 1 |
QY 301 GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
DB 1 |
QY 320 GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 379
DB 1 |
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATGAGAAATGTTCACTCAGAGATGACAGTTTG 420
DB 1 |
QY 380 CGCTGTTTCTATGGCTTTCAAAATCTCATGAGAAATGTTCACTCAGAGATGACAGTTTG 439
DB 1 |
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATGAA 480
DB 1 |
QY 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATGAA 499
DB 1 |
QY 481 ACCATGCCCTATGTTAAGAAAGAGCAGATTTGGGCTTTGGATGATAGCAGATAGAAA 540
DB 1 |
QY 500 ACCATGCCCTATGTTAAGAAAGAGCAGATTTGGGCTTTGGATGATAGCAGATAGAAA 559
DB 1 |
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAAGAGGATTTTCTTCTCAGGA 600
DB 1 |
QY 560 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAAGAGGATTTTCTTCTCAGGA 619
DB 1 |
QY 601 TCTTTGCTGCTATATCTGGCTAAAGAGAGGCTTTATGCCAGACTCACTTTTTC 660
DB 1 |
QY 620 TCTTTGCTGCTATATCTGGCTAAAGAGAGGCTTTATGCCAGACTCACTTTTTC 679
DB 1 |
QY 661 AATGAACCTCATCAGCAGATGAAGACTTCTACCTGCTGCTTTGCTGCTGATGTTCCAA 720
DB 1 |
QY 680 AATGAACCTCATCAGCAGATGAAGACTTCTACCTGCTGCTTTGCTGCTGATGTTCCAA 739
DB 1 |
QY 721 TACTTAGTAAATAGCCCTTCAGAAAGAGGCTCAGGAGATCAATTTGATGCTGTCAA 780
DB 1 |
QY 740 TACTTAGTAAATAGCCCTTCAGAAAGAGGCTCAGGAGATCAATTTGATGCTGTCAA 799
DB 1 |
QY 781 ATTGAGCAGAGTTTAAACAGAGCCCTTGCAGTGGCCCTCATTTGGAATGAATTCATT 840
DB 1 |
QY 800 ATTGAGCAGAGTTTAAACAGAGCCCTTGCAGTGGCCCTCATTTGGAATGAATTCATT 859
DB 1 |
QY 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTTACTTTGTGGAACTTGGATTCTCA 900
DB 1 |

Db 860 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTTACTTTGTGGAACTTGGATTCTCA 919
QY 901 AAGGTTTTTTCAGGCGAGAAATCCCTTTTCATTTTATGGAACAACTTTCTTTAGAGGAAAA 960
DB 920 AAGGTTTTTTCAGGCGAGAAATCCCTTTTCATTTTATGGAACAACTTTCTTTAGAGGAAAA 979
QY 961 ACAAAATTTCTTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTTTCAGTTATGCGAGAAACC 1020
DB 980 ACAAAATTTCTTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTTTCAGTTATGCGAGAAACC 1039
QY 1021 ACAGATAACGCTTTCACCTTCGATGCGAGATTTT 1053
DB 1040 ACAGATAACGCTTTCACCTTCGATGCGAGATTTT 1072

RESULT 14
BD064766
LOCUS Novel protein and DNA thereof.
DEFINITION BD064766
ACCESSION BD064766
VERSION BD064766.1 GI:22610369
KEYWORDS JP 2001269184-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1081)
AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE Novel protein and DNA thereof
JOURNAL Patent: JP 2001269184-A 3 02-OCT-2001;
        YUSUKE NAKAMURA, TAKEEDA CHEMICAL INDUSTRIES LTD
COMMENT OS Homo sapiens (human)
        PN JP 2001269184-A/3
        PD 02-OCT-2001
        PI 27-JUN-2000 JP 2000192401
        PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
        C12N15/09, A61K31/711, A61K38/00, A61K38/43, A61K45/00, A61K48/00, PC
        A61P35/00,
        PC A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC
        , C12N5/10, C12N9/02,
        PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53// (C12N15/09,
        C12R1:91),
        PC (C12N15/09, C12R1:19), (C12N9/02, C12R1:91), (C12P21/02, C12R1:91),
        PC (C12P21/08, C12R1:91), (C12N15/00, A61K37/02, A61K37/48, C12N5/00,
        PC (C12N15/00, C12R1:91), (C12N15/00, C12R1:19)
        CC Novel protein and DNA thereof
        FH Key Location/Qualifiers
        FT source 1..1081
        FT /organism="Homo sapiens (human)".
FEATURES source
        1..1081
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"

ORIGIN
Query Match      99.8%; Score 1051.4; DB 6; Length 1081;
Best Local Similarity 99.9%; Pred. No. 3e-255;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCGGAAAGCGCGGAGCGGCTGGATCAGGATGAGATGAGATCATCTTCA 60
DB 20 ATGGGCGACCGGAAAGCGCGGAGCGGCTGGATCAGGATGAGATGAGATCATCTTCA 79
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGC 120
DB 20 ATGGGCGACCGGAAAGCGCGGAGCGGCTGGATCAGGATGAGATGAGATCATCTTCA 79
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGC 120
DB 80 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGC 139
QY 121 CGGTTTGTCATCTTTCCATCCAGTACCCCTGATATTTGGAAATGTATAACAGGACAG 180
DB 80 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGC 139
QY 121 CGGTTTGTCATCTTTCCATCCAGTACCCCTGATATTTGGAAATGTATAACAGGACAG 180
DB 121 CGGTTTGTCATCTTTCCATCCAGTACCCCTGATATTTGGAAATGTATAACAGGACAG 180
QY 140 CGGTTTGTCATCTTTCCATCCAGTACCCCTGATATTTGGAAATGTATAACAGGACAG 199
DB 140 CGGTTTGTCATCTTTCCATCCAGTACCCCTGATATTTGGAAATGTATAACAGGACAG 199
QY 181 GCTTCCTCTGGACAGAGAGGTCGATTAATCAAGAGTCTCCCTCACTGGAACAG 240
DB 181 GCTTCCTCTGGACAGAGAGGTCGATTAATCAAGAGTCTCCCTCACTGGAACAG 240
QY 200 GCTTCCTCTGGACAGAGAGGTTGACTTATCAAGAGTCTCCCTCACTGGAACAG 259
DB 200 GCTTCCTCTGGACAGAGAGGTTGACTTATCAAGAGTCTCCCTCACTGGAACAG 259
QY 241 CTTAAAGCAGATGAGAGTACTTCTCTCTCACATCTTAGCTTTTGGACGAGGAT 300
DB 241 CTTAAAGCAGATGAGAGTACTTCTCTCTCACATCTTAGCTTTTGGACGAGGAT 300
QY 260 CTTAAAGCAGATGAGAGTACTTCTCTCTCACATCTTAGCTTTTGGACGAGGAT 319
DB 260 CTTAAAGCAGATGAGAGTACTTCTCTCTCACATCTTAGCTTTTGGACGAGGAT 319
QY 301 GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
DB 301 GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
QY 320 GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 379
DB 320 GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 379
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATGAGAAATGTTCACTCAGAGATGACAGTTTG 420
DB 361 CGCTGTTTCTATGGCTTTCAAAATCTCATGAGAAATGTTCACTCAGAGATGACAGTTTG 420
QY 380 CGCTGTTTCTATGGCTTTCAAAATCTCATGAGAAATGTTCACTCAGAGATGACAGTTTG 439
DB 380 CGCTGTTTCTATGGCTTTCAAAATCTCATGAGAAATGTTCACTCAGAGATGACAGTTTG 439
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATGAA 480
DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATGAA 480
QY 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATGAA 499
DB 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATGAA 499
QY 481 ACCATGCCCTATGTTAAGAAAGAGCAGATTTGGGCTTTGGATGATAGCAGATAGAAA 540
DB 481 ACCATGCCCTATGTTAAGAAAGAGCAGATTTGGGCTTTGGATGATAGCAGATAGAAA 540
QY 500 ACCATGCCCTATGTTAAGAAAGAGCAGATTTGGGCTTTGGATGATAGCAGATAGAAA 559
DB 500 ACCATGCCCTATGTTAAGAAAGAGCAGATTTGGGCTTTGGATGATAGCAGATAGAAA 559
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAAGAGGATTTTCTTCTCAGGA 600
DB 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAAGAGGATTTTCTTCTCAGGA 600
QY 560 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAAGAGGATTTTCTTCTCAGGA 619
DB 560 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAAGAGGATTTTCTTCTCAGGA 619
QY 601 TCTTTGCTGCTATATCTGGCTAAAGAGAGGCTTTATGCCAGACTCACTTTTTC 660
DB 601 TCTTTGCTGCTATATCTGGCTAAAGAGAGGCTTTATGCCAGACTCACTTTTTC 660
QY 620 TCTTTGCTGCTATATCTGGCTAAAGAGAGGCTTTATGCCAGACTCACTTTTTC 679
DB 620 TCTTTGCTGCTATATCTGGCTAAAGAGAGGCTTTATGCCAGACTCACTTTTTC 679
QY 661 AATGAACCTCATCAGCAGATGAAGACTTCTACCTGCTGCTTTGCTGCTGATGTTCCAA 720
DB 661 AATGAACCTCATCAGCAGATGAAGACTTCTACCTGCTGCTTTGCTGCTGATGTTCCAA 720
QY 680 AATGAACCTCATCAGCAGATGAAGACTTCTACCTGCTGCTTTGCTGCTGATGTTCCAA 739
DB 680 AATGAACCTCATCAGCAGATGAAGACTTCTACCTGCTGCTTTGCTGCTGATGTTCCAA 739
QY 721 TACTTAGTAAATAGCCCTTCAGAAAGAGGCTCAGGAGATCAATTTGATGCTGTCAA 780
DB 721 TACTTAGTAAATAGCCCTTCAGAAAGAGGCTCAGGAGATCAATTTGATGCTGTCAA 780
QY 740 TACTTAGTAAATAGCCCTTCAGAAAGAGGCTCAGGAGATCAATTTGATGCTGTCAA 799
DB 740 TACTTAGTAAATAGCCCTTCAGAAAGAGGCTCAGGAGATCAATTTGATGCTGTCAA 799
QY 781 ATTGAGCAGAGTTTAAACAGAGCCCTTGCAGTGGCCCTCATTTGGAATGAATTCATT 840
DB 781 ATTGAGCAGAGTTTAAACAGAGCCCTTGCAGTGGCCCTCATTTGGAATGAATTCATT 840
QY 800 ATTGAGCAGAGTTTAAACAGAGCCCTTGCAGTGGCCCTCATTTGGAATGAATTCATT 859
DB 800 ATTGAGCAGAGTTTAAACAGAGCCCTTGCAGTGGCCCTCATTTGGAATGAATTCATT 859
QY 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTTACTTTGTGGAACTTGGATTCTCA 900
DB 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTTACTTTGTGGAACTTGGATTCTCA 900
```

Qy	181	GCTTCTCTCTGGACAGCAGAGAGGTGCGACTTATCAAGAGTATCTCCCTCACTGGAACAAG	240
Ds	200	GCTTCTCTCTGGACAGCAGAGAGAGGTGCGACTTATCAAGAGTATCTCCCTCACTGGAACAAG	259
Qy	241	CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGGCTTTTGGACGAGTGAT	300
Ds	260	CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGGCTTTTGGACGAGTGAT	319
Qy	301	GGAAATGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGTGCGAGGTCTCCAGAGGCT	360
Ds	320	GGAAATGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGTGCGAGGTCTCCAGAGGCT	379
Qy	361	CGCTGTTTCTATGCTCTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGACAGTTTG	420
Ds	380	CGCTGTTTCTATGCTCTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGACAGTTTG	439
Qy	421	CTGATAGACATTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA	480
Ds	440	CTGATAGACATTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA	499
Qy	481	ACCATGCCCTATGTTAAGAAAGAGCAGATTTGGGCTTGGATGATGACAGATAGAAA	540
Ds	500	ACCATGCCCTATGTTAAGAAAGAGCAGATTTGGGCTTGGATGATGACAGATAGAAA	559
Qy	541	TCTACTTTTGGGAAAGAGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA	600
Ds	560	TCTACTTTTGGGAAAGAGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA	619
Qy	601	TCTTTTGTGCTATATTTCTGCTTAAAGAGAGAGTCTTATGCGCAGACACTCTTTTTC	660
Ds	620	TCTTTTGTGCTATATTTCTGCTTAAAGAGAGAGTCTTATGCGCAGACACTCTTTTTC	679
Qy	661	AATGAACCTCATCAGCAGATGAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA	720
Ds	680	AATGAACCTCATCAGCAGATGAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA	739
Qy	721	TACTTAGTAAATAGCCCTCAGAGAAAGGCTCAGGAGAGTCACTTTGTGCTGTCAAA	780
Ds	740	TACTTAGTAAATAGCCCTCAGAGAAAGGCTCAGGAGAGTCACTTTGTGCTGTCAAA	799
Qy	781	ATTGAGCAGGAGTTTAAACAGAGCCCTGCGAGTGGCTCAITGGAAATGAAITGCAIT	840
Ds	800	ATTGAGCAGGAGTTTAAACAGAGCCCTGCGAGTGGCTCAITGGAAATGAAITGCAIT	859
Qy	841	TTGATGAAACAGTACATTTGATGCTGACAGATTACTTTGGAACTTGGATTTCTCA	900
Ds	860	TTGATGAAACAGTACATTTGATGCTGACAGATTACTTTGGAACTTGGATTTCTCA	919
Qy	901	AAGGTTTTTTCAGGCAGAGAAATCCCTTTGATTTTATGAAACATTTCTTTAGAGGAAA	960
Ds	920	AAGGTTTTTTCAGGCAGAGAAATCCCTTTGATTTTATGAAACATTTCTTTAGAGGAAA	979
Qy	961	ACAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC	1020
Ds	980	ACAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC	1039
Qy	1021	ACAGATAACGCTTTCACCTTGGATGCGAGATTTT	1053
Ds	1040	ACAGATAACGCTTTCACCTTGGATGCGAGATTTT	1072
RESULT 15			
BD093079			
LOCUS			
DEFINITION			
Novel protein and its DNA.			
ACCESSION			
BD093079			
VERSION			
BD093079.1			
KEYWORDS			
WO 0100799-A/3.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1 (bases 1 to 1081)			

AUTHORS		Nakamura, Y., Arakawa, H. and Tanaka, H.	
TITLE		Novel protein and its DNA	
JOURNAL		Patent: WO 0100799-A 3 04-JAN-2001;	
		TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,	
		HIROSHI TANAKA	
COMMENT		OS Homo sapiens (human)	
		PN WO 0100799-A/3	
		PD 04-JAN-2001	
		PF 27-JUN-2000 WO 2000JP004189	
		PR 28-JUN-1999 JP 99P 181131, 06-JUL-1999 JP 99P 192391 PR	
		PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC	
		C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N5/10 PC	
		C12P21/02, A61K38/44,	
		PC A61K48/00, C07K16/40, G01N33/53, C12Q1/26	
		CC	
FEATURES		FH Key Location/Qualifiers.	
Source		1. .1081	
		/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
ORIGIN			
Query Match		99.8%; Score 1051.4; DB 6; Length 1081;	
Best Local Similarity		99.9%; Pred. No. 3e-255;	
Matches 1052; Conservative		0; Mismatches 1; Indels 0; Gaps 0;	
Qy	1	ATGGCGGACCCGGAAAGGCGCGGCGTGGATCAGGATGAGAGATCATCTTCA	60
Ds	20	ATGGCGGACCCGGAAAGGCGCGGCGTGGATCAGGATGAGAGATCATCTTCA	79
Qy	61	GACACCAACGAAAGTGAATTAAGTCAATGAAGGCCACTCTTAAGAAAGATTCTCGC	120
Ds	80	GACACCAACGAAAGTGAATTAAGTCAATGAAGGCCACTCTTAAGAAAGATTCTCGC	139
Qy	121	CGGTTTGTCTATCTTTTCAATCCAGTACCTCTGATATTTGGAAATCTATAAAGCAGCACAG	180
Ds	140	CGGTTTGTCTATCTTTTCAATCCAGTACCTCTGATATTTGGAAATCTATAAAGCAGCACAG	199
Qy	181	GCTTCTCTTCTGGACAGCAGAGAGGTGAGATTCAGATTCCTCCCTCACTGGAAACAAG	240
Ds	200	GCTTCTCTTCTGGACAGCAGAGAGGTGAGATTCAGATTCCTCCCTCACTGGAAACAAG	259
Qy	241	CTTAAACAGATGAGAAGTACTTCACTCTCTCACATCTTAGGCTTTTTCAGGCGAGTGAT	300
Ds	260	CTTAAACAGATGAGAAGTACTTCACTCTCTCACATCTTAGGCTTTTTCAGGCGAGTGAT	319
Qy	301	GGAAATGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGGTTCAGAGGCT	360
Ds	320	GGAAATGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGGTTCAGAGGCT	379
Qy	361	CGCTGTTTCTATGCGCTTTCAAATTTCTCATCAGAGATGTTTCACTCAGAGATGTACAGTTTG	420
Ds	380	CGCTGTTTCTATGCGCTTTCAAATTTCTCATCAGAGATGTTTCACTCAGAGATGTACAGTTTG	439
Qy	421	CTGATAGACATTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA	480
Ds	440	CTGATAGACATTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA	499
Qy	481	ACCATGCCCTATGTTAAGAAAGAGCAGATTTGGGCTTGGATGATGACAGATAGAAA	540
Ds	500	ACCATGCCCTATGTTAAGAAAGAGCAGATTTGGGCTTGGATGATGACAGATAGAAA	559
Qy	541	TCTACTTTTGGGAAAGAGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA	600
Ds	560	TCTACTTTTGGGAAAGAGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA	619
Qy	601	TCTTTTGTGCTATATTTCTGCTTAAAGAGAGAGGTCTTATGCGCAGACACTCTTTTTC	660
Ds	620	TCTTTTGTGCTATATTTCTGCTTAAAGAGAGAGGTCTTATGCGCAGACACTCTTTTTC	679
Qy	661	AATGAACCTCATCAGCAGATGAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA	720

[illegible]

Search completed: September 26, 2005, 01:35:17
Job time : 3285 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2005, 18:03:30 ; Search time 463 Seconds
(without alignments)
13463.256 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1053

Sequence: 1 atggcgaccgcgaaggcc.....tcaccttgatgcagatttt 1053

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_16Dec04.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1053	100.0	1053	4	Aaf32438 Human rib
2	1053	100.0	2596	3	Aa12411 CDNA enco
3	1053	100.0	4955	4	Aaf32439 Human rib
4	1053	100.0	4955	13	ADR24210 Breast ca
5	1051.4	99.8	1053	4	Aaf32447 Human rib
6	1051.4	99.8	1081	4	Aaf32440 Human rib
7	1049.8	99.7	1601	4	Aa14924 Human cDN
8	650.6	61.8	706	13	Adq57092 Novel can
9	578.4	54.9	1989	3	Aac78111 Human can
10	578.4	54.9	2215	12	Adk70302 Respirato
11	578.4	54.9	2216	10	Adj56536 Murine cD
12	578.4	54.9	2482	4	Aa544917 Human con
13	578.4	54.9	2500	6	ABL65414 Lung canc
14	578.4	54.9	2500	6	ABL65517 Lung canc
15	578.4	54.9	2500	6	ABL65859 Lung canc
16	578.4	54.9	2500	8	ABX10335 DNA enco
17	578.4	54.9	2500	11	ADI32044 Human cDN
18	578.4	54.9	2500	12	ADN04443 Antipsoi
19	578.4	54.9	2500	12	Adq09273 Human RRM
20	578.4	54.9	2500	13	ACN37637 Tumour-as

21	578.4	54.9	3294	4	AAH73225	Aah73225 Human cer
22	578.4	54.9	3393	12	ADN03788	Adn03788 Antipsoi
23	578.4	54.9	3393	12	ADO19225	Ado19225 Human PRO
24	569	54.0	2641	4	AAS44745	Aas44745 Human ful
25	568.8	54.0	1170	2	AAV05641	Aav05641 Human rib
26	561	53.3	1328	12	ADO57308	Ado57308 DNA enco
27	547.8	52.0	978	4	AAS41006	Aas41006 cDNA enco
28	545.4	51.8	977	6	ABL90228	Ab190228 Human pol
29	528.8	50.2	608	4	AAH07707	Aah07707 Human cDN
30	514	48.8	1371	5	AAS79474	Aas79474 DNA enco
31	463	44.0	1289	4	ABL14627	Ab114627 Drosophil
32	442	42.0	1146	6	AAS62387	Aas62387 cDNA sequ
33	441	41.9	481	2	AAX51874	Aax51874 Human sec
34	432.2	41.0	1218	8	ABT20895	Abt20895 Aspergill
c	429.8	40.8	14176	2	AAT84564	Aat84564 Swinepox
36	429.8	40.8	14176	4	AAF84949	Aaf84949 Reverse c
37	420.4	39.9	3945	4	ABL14626	Aaf84948 Nucleotid
38	420.4	39.9	3945	4	ABL14626	Ab114626 Drosophil
39	403.2	38.3	963	13	ADS58378	AdS58378 Bacteri
40	397	37.7	1146	8	ABT19075	Abt19075 Aspergill
41	392.6	37.3	1173	13	ADT47814	Adt47814 Bacteri
42	392	37.2	1242	6	ABZ32250	Abz32250 Candida a
43	386.6	36.7	1292	13	ADS49379	AdS49379 Bacteri
44	380.4	36.1	1200	10	ACC61238	Acc61238 Gene sequ
45	380.4	36.1	1200	10	ADK63241	Adk63241 Disease t

ALIGNMENTS

RESULT 1

AAF32438
ID AAF32438 standard; cDNA; 1053 BP.

AC AAF32438;

DT 18-APR-2001 (first entry)

DE Human ribonucleotide reductase TP53R2H nucleotide sequence SEQ ID NO:2.

KW Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.

XX Homo sapiens.

PN WO200100799-A1.

PD 04-JAN-2001.

PF 27-JUN-2000; 2000WO-JP004189.

PR 28-JUN-1999; 99JP-00181131.

PR 06-JUL-1999; 99JP-00192391.

PR 21-JAN-2000; 2000JP-00017770.

XX (TAKE) TAKEDA CHEM IND LTD.

PA (NAKA/) NAKAMURA Y.

XX Nakamura Y, Arakawa H, Tanaka H;

DR WPI; 2001-112446/12.

DR P-PSDB; AAB69050.

XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for diagnosis, treatment and prevention of cancer.

PS Claim 5; Fig 1-3; 102pp; Japanese.

XX The present sequence encodes a human ribonucleotide reductase designated TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of the DNA repair mechanism and its activity is induced by p53. It can be used for the treatment, prevention and diagnosis of a wide range of cancers

SQ Sequence 1053 BP; 314 A; 192 C; 246 G; 301 T; 0 U; 0 Other;

Query Match		100.0%;	Score 1053;	DB 4;	Length 1053;
Best Local Similarity		100.0%;	Pred. No. 1e-290;		
Matches 1053;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGGCGCACC	GGGAAAGCGCGG	AGCGCGGGCTGGATCAGGATGAGATCATCTTCA	60
Db	1	ATGGCGCACC	GGGAAAGCGCGG	AGCGCGGGCTGGATCAGGATGAGATCATCTTCA	60
Qy	61	GACACCAACG	AAAGTAAAGTCA	AAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC	120
Db	61	GACACCAACG	AAAGTAAAGTCA	AAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC	120
Qy	121	CGGTTGCTCAT	CTTTCCAAATCCAGT	ACCCTGATATTTGGAAATGTATAAACAGGCACAG	180
Db	121	CGGTTGCTCAT	CTTTCCAAATCCAGT	ACCCTGATATTTGGAAATGTATAAACAGGCACAG	180
Qy	181	GCTTCCTTCTG	GACAGAGAGGTG	AGCTTATCAAAAGATCTCCCTCACTGGAACAAG	240
Db	181	GCTTCCTTCTG	GACAGAGAGGTG	AGCTTATCAAAAGATCTCCCTCACTGGAACAAG	240
Qy	241	CTTAAAGCAG	ATGAGTAGTACT	CTCATCTCTCACATCTTAGCCTTTTTCGAGCCAGTGAT	300
Db	241	CTTAAAGCAG	ATGAGTAGTACT	CTCATCTCTCACATCTTAGCCTTTTTCGAGCCAGTGAT	300
Qy	301	GGAATTGTAAT	CAAAATTTGGTGG	AGCGCTTTTAGTCAGGAGGTGACAGGCTTCCAGAGGCT	360
Db	301	GGAATTGTAAT	CAAAATTTGGTGG	AGCGCTTTTAGTCAGGAGGTGACAGGCTTCCAGAGGCT	360
Qy	361	CGCTGTTTCTA	TGCTTCAAAATCT	CATCAGAGATGTTTATTTATGCAATTTGAA	420
Db	361	CGCTGTTTCTA	TGCTTCAAAATCT	CATCAGAGATGTTTATTTATGCAATTTGAA	420
Qy	421	CTGATACACAT	TTACATCAGAGT	CCCCAAGAAAGGGAATTTTATTTATGCAATTTGAA	480
Db	421	CTGATACACAT	TTACATCAGAGT	CCCCAAGAAAGGGAATTTTATTTATGCAATTTGAA	480
Qy	481	ACCATGCCCT	TATGTTAAGAAAAA	AGCAGATTTGGGCCCTTCGATGATGAGATAGAAAA	540
Db	481	ACCATGCCCT	TATGTTAAGAAAAA	AGCAGATTTGGGCCCTTCGATGATGAGATAGAAAA	540
Qy	541	TCTACTTTTGG	GAAGAGTGGTGG	CCCTTTGCTGTAGAGAGGATTTTCTCTCAGGA	600
Db	541	TCTACTTTTGG	GAAGAGTGGTGG	CCCTTTGCTGTAGAGAGGATTTTCTCTCAGGA	600
Qy	601	TCCTTTGCTGT	ATATCTGCTTAA	AGAGAGGCTTTATGCCAGCACTCTTTTTC	660
Db	601	TCCTTTGCTGT	ATATCTGCTTAA	AGAGAGGCTTTATGCCAGCACTCTTTTTC	660
Qy	661	AATGAACCTCA	TACAGAGATGAAG	AGCTTCTGCTGTAGAGAGGATTTTCTCTCAGGA	720
Db	661	AATGAACCTCA	TACAGAGATGAAG	AGCTTCTGCTGTAGAGAGGATTTTCTCTCAGGA	720
Qy	721	TACTTAGTAA	TAATAGCCTTCAG	AGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAAA	780
Db	721	TACTTAGTAA	TAATAGCCTTCAG	AGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAAA	780
Qy	781	ATTGAGCAGAG	TTTTTAAAGAGC	CTTCCAGTTGGCCCTCATTTGGAATGATTCATT	840
Db	781	ATTGAGCAGAG	TTTTTAAAGAGC	CTTCCAGTTGGCCCTCATTTGGAATGATTCATT	840
Qy	841	TTGATGAAA	CAGTACATTTG	ATGAGTGTAGCTGACAGATTCTTTGGAACCTTGAATTCCTCA	900
Db	841	TTGATGAAA	CAGTACATTTG	ATGAGTGTAGCTGACAGATTCTTTGGAACCTTGAATTCCTCA	900
Qy	901	AAGGTTTTTTC	AGGAGAAAATCCT	TTTTGATTTTATGGAATAACATTTCTTTAGAGGAAAA	960
Db	901	AAGGTTTTTTC	AGGAGAAAATCCT	TTTTGATTTTATGGAATAACATTTCTTTAGAGGAAAA	960
Qy	961	ACAAATTTCT	TTGAGAAAACAG	TTTTTCAGATATCAGCGTTTTTGCGATTTATGGCAGAAACC	1020
Db	961	ACAAATTTCT	TTGAGAAAACAG	TTTTTCAGATATCAGCGTTTTTGCGATTTATGGCAGAAACC	1020

Qy	1021	ACAGATAACG	CTTTTACCTTGC	ATGATGATGATTTT	1053
Db	1021	ACAGATAACG	CTTTTACCTTGC	ATGATGATGATTTT	1053

RESULT 2
ID AAA12411 standard; cDNA; 2596 BP.
XX AC AAA12411;
XX XX 25-JUL-2000 (first entry)
XX DE cDNA encoding a human RNA-associated protein.

XX Human; RNA-associated protein; cell proliferation; cancer; inflammation;
XX immune response; reproductive disorder; acinic keratosis;
XX KW atherosclerosis; arteriosclerosis; bursitis; cirrhosis; hepatitis;
XX KW mixed connective tissue disease; myelofibrosis; primary thrombocythemia;
XX KW paroxysmal nocturnal hemoglobinuria; polycythemia vera; psoriasis;
XX KW trauma; ss.

XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 56..1111
XX FT /*tag= a
XX FT /product= "RNA-associated protein"

XX W0200015799-A2.
XX FN
XX PD 23-MAR-2000.

XX PF 17-SEP-1999; 99WO-US021688.
XX PR 17-SEP-1998; 98US-00156039.
XX PR 22-SEP-1998; 98US-00158720.
XX PR 04-NOV-1998; 98US-00186815.
XX PR 08-APR-1999; 99US-0128660P.
XX (INCY-) INCYTE PHARM INC.

XX Tang YT, Corley NC, Guegler KJ, Gorgone GA, Patterson C;
XX Hillman JL, Baughn MR, Lal P, Azimzai Y, Yue H, Yang J;
XX WPI; 2000-271437/23.
XX P-PSDB; AAY84439.

XX New polypeptides and polynucleotides, useful for preventing and treating
XX a disorder associated with increased or decreased expression of RNA
XX associated proteins.

XX Claim 9; Page 120-121; 131pp; English.

XX The present sequence encodes a human RNA-associated protein. The
XX expression of RNA-associated proteins is closely associated with
XX reproductive tissues, nervous tissues, cell proliferation including
XX cancer, inflammation and immune responses, and so they may be used for
XX diagnosis, treatment or prevention of cell proliferative,
XX immune/inflammatory disorders, and reproductive disorders. Diseases and
XX disorders which may be treated include acinic keratosis,
XX atherosclerosis, arteriosclerosis, bursitis, cirrhosis, hepatitis, mixed
XX connective tissue disease, myelofibrosis, paroxysmal nocturnal
XX hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia
XX and cancers, and trauma

XX Sequence 2596 BP; 783 A; 418 C; 510 G; 885 T; 0 U; 0 Other;

Query Match 100.0%; Score 1053; DB 3; Length 2596;
Best Local Similarity 100.0%; Pred. No. 1.5e-290;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGCACC

```
Db 56 ATGGGCGACCCGGAAGCCGGAAGCCGGCTGGATCAGGATGAGAGATCATCTTCA 115
Qy 61 GACACCAACGAAAGTGAATAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Db 116 GACACCAACGAAAGTGAATAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 175
Qy 121 CGTTTGTGTCATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATAAACAGCACAG 180
Db 176 CGTTTGTGTCATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATAAACAGCACAG 235
Qy 181 GCTTCCTTCTGGACAGCAGAGAGGTGCACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 240
Db 236 GCTTCCTTCTGGACAGCAGAGAGGTGCACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 295
Qy 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTTAGCCTTTTGGCAGCCAGTGAT 300
Db 296 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTTAGCCTTTTGGCAGCCAGTGAT 355
Qy 301 GGAATTGTAAATGAAAAATTTGGTGGAGCGCTTTAGTCAAGAGGTGCGAGTTCCAGAGGCT 360
Db 356 GGAATTGTAAATGAAAAATTTGGTGGAGCGCTTTAGTCAAGAGGTGCGAGTTCCAGAGGCT 415
Qy 361 CGCTGTTCTATGGCTTTCAAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 416 CGCTGTTCTATGGCTTTCAAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 475
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 480
Db 476 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 535
Qy 481 ACCATGCCCTATGTTAAGAAAAAGCAGATGGGCTTGGCATGATAGCAGATAGAAAA 540
Db 536 ACCATGCCCTATGTTAAGAAAAAGCAGATGGGCTTGGCATGATAGCAGATAGAAAA 595
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTCTGCTGTAGAGAGTGTCTTCTTCAGGA 600
Db 596 TCTACTTTTGGGAAAGAGTGGTGGCTTCTGCTGTAGAGAGTGTCTTCTTCAGGA 655
Qy 601 TCTTTTGTGCTATATTTCTGCTTAAAGAGAGAGTCTTATGCCAGGACTCATTCTTCC 660
Db 656 TCTTTTGTGCTATATTTCTGCTTAAAGAGAGAGTCTTATGCCAGGACTCATTCTTCC 715
Qy 661 AATGAACCTCATCAGCAGATGAGGACTTCACATGTGACTTTGCTTGGCTGATGTCCAA 720
Db 716 AATGAACCTCATCAGCAGATGAGGACTTCACATGTGACTTTGCTTGGCTGATGTCCAA 775
Qy 721 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAA 780
Db 776 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAA 835
Qy 781 ATTGAGCGAGGAGTTTAAACAGAGCCCTTGGCCAGTTGGCCCTCATTTGGAATGAATTGCA 840
Db 836 ATTGAGCGAGGAGTTTAAACAGAGCCCTTGGCCAGTTGGCCCTCATTTGGAATGAATTGCA 895
Qy 841 TTGATGAACAGTACATTGAGTTGTAGTACACAGATTTACTTGTGGAATTTGGATTTCTCA 900
Db 896 TTGATGAACAGTACATTGAGTTGTAGTACACAGATTTACTTGTGGAATTTGGATTTCTCA 955
Qy 901 AAGGTTTTTTCAGGCAGAAAAATCCTTTTGAATTTTATGGAACATTTCTTTAGAAAGGAAA 960
Db 956 AAGGTTTTTTCAGGCAGAAAAATCCTTTTGAATTTTATGGAACATTTCTTTAGAAAGGAAA 1015
Qy 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTGCAGTTATGGCAGAAAC 1020
Db 1016 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTGCAGTTATGGCAGAAAC 1075
Qy 1021 ACAGATAACGCTTTCACCTTGGATGCGAGTTTTT 1053
Db 1076 ACAGATAACGCTTTCACCTTGGATGCGAGTTTTT 1108
```

RESULT 3

AAF32439

ID AAF32439 standard; cDNA; 4955 BP.

XX AC AAF32439;

XX DT 18-APR-2001 (first entry)

XX DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:3.

XX KW Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.

XX OS Homo sapiens.

XX PN WO200100799-A1.

XX PD 04-JAN-2001.

XX PF 27-JUN-2000; 2000WO-JP004189.

XX PR 28-JUN-1999; 99JP-00181131.

XX PR 06-JUL-1999; 99JP-00192391.

XX PR 21-JAN-2000; 2000JP-00017770.

XX PA (TAKB) TAKEDA CHEM IND LTD.

XX PA (NAKA/) NAKAMURA Y.

XX PI Nakamura Y, Arakawa H, Tanaka H;

XX DR WPI; 2001-112446/12.

XX PT Ribonucleotide reductase involved in DNA repair and DNA encoding it, for

XX PS diagnosis, treatment and prevention of cancer.

XX PS Example 2; Page 87-90; 102pp; Japanese.

XX The present invention describes a human ribonucleotide reductase designated TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of the DNA repair mechanism and its activity is induced by p53. It can be used for the treatment, prevention and diagnosis of a wide range of cancers. The present sequence represents a human ribonucleotide reductase related sequence which is used in an example from the present invention

XX SQ Sequence 4955 BP; 1526 A; 803 C; 999 G; 1627 T; 0 U; 0 Other;

Query Match 100.0%; Score 1053; DB 4; Length 4955;

Best Local Similarity 100.0%; Pred. No. 2.1e-290;

Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGGCGACCCGGAAGCCGGAAGCCGGCTGGATCAGGATGAGAGATCATCTTCA 60

Db 245 ATGGGCGACCCGGAAGCCGGAAGCCGGCTGGATCAGGATGAGAGATCATCTTCA 304

Qy 61 GACACCAACGAAAGTGAATAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120

Db 305 GACACCAACGAAAGTGAATAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 364

Qy 121 CGTTTGTGTCATCTTTCAAATCCAGTACCTGATTTTGGAAATGTATAAACAGGCACAG 180

Db 365 CGTTTGTGTCATCTTTCAAATCCAGTACCTGATTTTGGAAATGTATAAACAGGCACAG 424

Qy 181 GCTTCCTTCTGGACAGCAGAGAGGTGCACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 240

Db 425 GCTTCCTTCTGGACAGCAGAGAGGTGCACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 484

Qy 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCTCACTCTTAGCCCTTTTTCAGCAGCAGTGAT 300

Db 485 CTTAAAGCAGATGAGAAGTACTTCACTCTCTCACTCTTAGCCCTTTTTCAGCAGCAGTGAT 544

Qy 301 GGAATTGTAAATGAAAAATTTGGTGGAGCGCTTTAGTCAAGAGGTGCGAGTTCCAGAGGCT 360

Db 545 GGAATTGTAAATGAAAAATTTGGTGGAGCGCTTTAGTCAAGAGGTGCGAGTTCCAGAGGCT 604

Qy 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGATGTTTCACTCAGAGATGTACAGTTTG 420

Db 605 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 664
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 724
Qy 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCCCTTGCGATGCGATAGCAGATAGAAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCCCTTGCGATGCGATAGCAGATAGAAAA 784
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCCCTTGCTGCTGTGAGAGAGTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGGTGGCCCTTGCTGCTGTGAGAGAGTTTCTTCTCAGGA 844
Qy 601 TCTTTTGTCTGCTATTTCTGCTTAAGAAAGAGAGTCTTATGCCAGACACTCCTTTTTC 660
Db 845 TCTTTTGTCTGCTATTTCTGCTTAAGAAAGAGAGTCTTATGCCAGACACTCCTTTTTC 904
Qy 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 964
Qy 721 TACTTAGTAATAAGCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA 780
Db 965 TACTTAGTAATAAGCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA 1024
Qy 781 ATTGAGCAGAGTTTTTAAACAGAGCCCTTGCCAGTTGGCCCTCATTTGGAATGATTCATT 840
Db 1025 ATTGAGCAGAGTTTTTAAACAGAGCCCTTGCCAGTTGGCCCTCATTTGGAATGATTCATT 1084
Qy 841 TTGATGAAAAAGTACATGAGTTTGTAGCTGACAGATTAATTTGTGGAACCTTGGATTTCTCA 900
Db 1085 TTGATGAAAAAGTACATGAGTTTGTAGCTGACAGATTAATTTGTGGAACCTTGGATTTCTCA 1144
Qy 901 AAGTTTTTTCAGGAGAAAAATCCTTTTGATTTTATGAAAAACATTTCTTTAGAAAGGAAAA 960
Db 1145 AAGTTTTTTCAGGAGAAAAATCCTTTTGATTTTATGAAAAACATTTCTTTAGAAAGGAAAA 1204
Qy 961 ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGGGTTTTGTCAGTTATGGCAGAAACC 1020
Db 1205 ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGGGTTTTGTCAGTTATGGCAGAAACC 1264
Qy 1021 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1053
Db 1265 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1297

RESULT 4
ID ADR24210
XX ADR24210 standard; DNA; 4955 BP.
AC ADR24210;
XX
XX
DT 21-OCT-2004 (first entry)
XX
XX Breast cancer prognosis marker #71.
XX ds; breast cancer; prognosis; gene expression; diagnosis.
XX Homo sapiens.
XX W02004065545-A2.
XX
PD 05-AUG-2004.
XX
XX 15-JAN-2004; 2004WO-US001100.
XX
XX 15-JAN-2003; 2003US-00342887.
XX (ROSE-) ROSETTA INPHARMATICS LLC.
XX (NECA-) NETHERLANDS CANCER INST.
XX

PI Van't Veer LJ, He Y;
XX WPI; 2004-593473/57.
XX
XX Classifying a breast cancer patient according to prognosis comprises
PT determining the similarity between the level of expression of each of
PT five genes in a cell sample taken from patient, to control levels.
XX
XX Disclosure; SEQ ID NO 71; 226pp; English.
XX
XX The invention relates to a method of classifying a breast cancer patient
CC according to prognosis by determining the similarity between the level of
CC expression of each of five genes for which markers are listed in the
CC specification, in a cell sample taken from the breast cancer patient, to
CC control levels of expression for each respective five genes to obtain a
CC patient similarity value. The methods are useful for classifying a breast
CC cancer patient according to prognosis. Kits and computer program products
CC are useful for data analysis using the diagnostic, prognostic and
CC statistical methods of the invention. This sequence corresponds to a
XX marker used in the method of the invention.
SQ Sequence 4955 BP; 1526 A; 803 C; 999 G; 1627 T; 0 U; 0 Other;

Query Match 100.0%; Score 1053; DB 13; Length 4955;
Best Local Similarity 100.0%; Pred. No. 2.1e-290;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGCACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 245 ATGGCGCACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304

Qy 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCCTTAAGAAAGAGTTCTCGC 120
Db 305 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCCTTAAGAAAGAGTTCTCGC 364

Qy 121 CGGTTTGTCTATCTTTTCAATCCAGTACCTCATATTTTGGAAAAATGATAAAGAGCAG 180
Db 365 CGGTTTGTCTATCTTTTCAATCCAGTACCTCATATTTTGGAAAAATGATAAAGAGCAG 424

Qy 181 GCTTCCTTCTGGACAGCAGAGAGGTGCACTTATCAAGAGATCTCCCTCACTGGAAACAG 240
Db 425 GCTTCCTTCTGGACAGCAGAGAGGTGCACTTATCAAGAGATCTCCCTCACTGGAAACAG 484

Qy 241 CTTAAAGCAGATGAGAAAGTACTTCTCATCTCTCAGATCTTTAGCTTCAGGAGGTCCAGAGGCT 300
Db 485 CTTAAAGCAGATGAGAAAGTACTTCTCATCTCTCAGATCTTTAGCTTCAGGAGGTCCAGAGGCT 544

Qy 301 GGAATTCGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAAGGAGGTCCAGAGGCT 360
Db 545 GGAATTCGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAAGGAGGTCCAGAGGCT 604

Qy 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGATGTTTCACTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGAAATGTTTCACTCAGAGATGTACAGTTTG 664

Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTGAA 724

Qy 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCCCTTGCGATGCGATAGCAGATAGAAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCCCTTGCGATGCGATAGCAGATAGAAAA 784

Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCCCTTGCTGCTGTGAGAGAGTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGGTGGCCCTTGCTGCTGTGAGAGAGTTTCTTCTCAGGA 844

Qy 601 TCTTTTGTCTGCTATTTCTGCTTAAGAAAGAGAGTCTTATGCCAGGACTCCTTTTTC 660
Db 845 TCTTTTGTCTGCTATTTCTGCTTAAGAAAGAGAGTCTTATGCCAGGACTCCTTTTTC 904

Qy 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 720

Db 905 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCTGATGTTCCAA 964
Qy 721 TACTTAGTAATAAGCCTTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db 965 TACTTAGTAATAAGCCTTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 1024
Qy 781 ATTGAGCAGGAGTTTAAACAGAGACCTTGCAGTTCGCTCATTTGGAATGAATGCAAT 840
Db 1025 ATTGAGCAGGAGTTTAAACAGAGACCTTGCAGTTCGCTCATTTGGAATGAATGCAAT 1084
Qy 841 TTGATGAACAGTACATTTGAGTTTGTAGCTCAGAGATTACTTTGGAACCTTTGATTTCTCA 900
Db 1085 TTGATGAACAGTACATTTGAGTTTGTAGCTCAGAGATTACTTTGGAACCTTTGATTTCTCA 1144
Qy 901 AAGTTTTTTCAGGCGAGAAATCCTTTTGAATTTATGGAACATTTCTTTAGAGGAA 960
Db 1145 AAGTTTTTTCAGGCGAGAAATCCTTTTGAATTTATGGAACATTTCTTTAGAGGAA 1204
Qy 961 ACAATTTCTTTGAGAACGAGTTTCAGAGATCAGCGTTTTCAGTTATGCGAGAAACC 1020
Db 1205 ACAATTTCTTTGAGAACGAGTTTCAGAGATCAGCGTTTTCAGTTATGCGAGAAACC 1264
Qy 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053
Db 1265 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1297

RESULT 5
AAF32447
ID AAF32447 standard; cDNA; 1053 BP.
XX AC AAF32447;
XX XX
XX 18-APR-2001 (first entry)
XX DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:12.
XX KW Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
XX OS Homo sapiens.
XX PN W0200100799-A1.
XX PD 04-JAN-2001.
XX PF 27-JUN-2000; 2000WO-JP004189.
XX PR 28-JUN-1999; 99JP-00181131.
XX PR 06-JUL-1999; 99JP-00192391.
XX PR 21-JAN-2000; 2000JP-00017770.
XX PA (TAKE) TAKEDA CHEM IND LTD.
XX PA (NAKA) NAKAMURA Y.
XX XX
XX Nakamura Y, Arakawa H, Tanaka H;
XX WPI; 2001-112446/12.
XX DR Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
XX PT diagnosis, treatment and prevention of cancer.
XX PS Claim 6; Page 96-97; 102pp; Japanese.
XX CC The present invention describes a human ribonucleotide reductase
XX CC designated TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of
XX CC the DNA repair mechanism and its activity is induced by p53. It can be
XX CC used for the treatment, prevention and diagnosis of a wide range of
XX CC cancers. The present sequence represents a human ribonucleotide reductase
XX CC related sequence which is given in the exemplification of the present
XX CC invention
XX XX
XX Sequence 1053 BP; 314 A; 191 C; 246 G; 302 T; 0 U; 0 Other;

Query Match 99.8%; Score 1051.4; DB 4; Length 1053;
Best Local Similarity 99.9%; Pred. No. 2.9e-290;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGGGCGACCGGAAAGCGCGGAGCGGCTGGATCAGATCAGAGATCATCTTCA 60
Db 1 ATGGGCGACCGGAAAGCGCGGAGCGGCTGGATCAGATCAGAGATCATCTTCA 60
Qy 61 GACACCAACGAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Qy 121 CGGTTTGTCACTTTTCCAATCCAGTACCTTGATAATTTGGAAATGTATAAACAGGCACAG 180
Db 121 CGGTTTGTCACTTTTCCAATCCAGTACCTTGATAATTTGGAAATGTATAAACAGGCACAG 180
Qy 181 GCTTCTTCTGGACAGCAGAGAGTGCATTTATCAAGGATCTCCTCATCTGGAAACAAG 240
Db 181 GCTTCTTCTGGACAGCAGAGAGTGCATTTATCAAGGATCTCCTCATCTGGAAACAAG 240
Qy 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACTTTAGCCTTTTTCAGCCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACTTTAGCCTTTTTCAGCCAGTGAT 300
Qy 301 GGAATTTGTAATGAAAATTTTGTGGAGCGCTTTAGTCAGGAGGTTCAGAGGCT 360
Db 301 GGAATTTGTAATGAAAATTTTGTGGAGCGCTTTAGTCAGGAGGTTCAGAGGCT 360
Qy 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAAGATGTTTCACTCAGAGATGATACAGTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAAGATGTTTCACTCAGAGATGATACAGTTG 420
Qy 421 CTGATAGACATTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTTGA 480
Db 421 CTGATAGACATTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTTGA 480
Qy 481 ACCATGCCCTATGTTAAGAAAAGAGATTTGGGCTTTGCGATGATAGCAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAGAGATTTGGGCTTTGCGATGATAGCAGATAGAAA 540
Qy 541 TCTACTTTTGGGAAAGAGTGTGCGCTTTGCTGTAGAGAGGATTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGTGCGCTTTGCTGTAGAGAGGATTTTCTTCTCAGGA 600
Qy 601 TCTTTTGTCTGTATATTTCTGCTAAGAGAGAGTCTTATGCCAGGACTCATCTTTTCC 660
Db 601 TCTTTTGTCTGTATATTTCTGCTAAGAGAGAGTCTTATGCCAGGACTCATCTTTTCC 660
Qy 661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGTATGTTCCAA 720
Db 661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGTATGTTCCAA 720
Qy 721 TACTTAGTAATAAGCCTTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db 721 TACTTAGTAATAAGCCTTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAGACCTTGCAGTTCGCTCATTTGGAATGAATGCAAT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGACCTTGCAGTTCGCTCATTTGGAATGAATGCAAT 840
Qy 841 TTGATGAACAGTACATTTGAGTTTGTAGCTCAGAGATTACTTTGGAACCTTTGATTTCTCA 900
Db 841 TTGATGAACAGTACATTTGAGTTTGTAGCTCAGAGATTACTTTGGAACCTTTGATTTCTCA 900
Qy 901 AAGTTTTTTCAGGCGAGAAATCCTTTTGAATTTATGGAACATTTCTTTAGAGGAA 960
Db 901 AAGTTTTTTCAGGCGAGAAATCCTTTTGAATTTATGGAACATTTCTTTAGAGGAA 960
Qy 961 ACAATTTCTTTGAGAACGAGTTTCAGAGATCAGCGTTTTCAGTTATGCGAGAAACC 1020
Db 961 ACAATTTCTTTGAGAACGAGTTTCAGAGATCAGCGTTTTCAGTTATGCGAGAAACC 1020
Qy 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053

```

Db      1021 ACAGATAACGTCCTTCCACCTTGGATGCGAGATTTT 1053
|||||
RESULT 6
AAF32440
ID      AAF32440 standard; cDNA; 1081 BP.
XX
AC      AAF32440;
XX
DT      18-APR-2001 (first entry)
XX
DE      Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:4.
XX
KW      Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
XX
OS      Homo sapiens.
XX
PN      WO200100799-A1.
XX
PD      04-JAN-2001.
XX
PF      27-JUN-2000; 2000WO-JP004189.
XX
PR      28-JUN-1999; 99JP-00181131.
XX
PR      06-JUL-1999; 99JP-00192391.
XX
PR      21-JAN-2000; 2000JP-00017770.
XX
PA      (TAKE ) TAKEDA CHEM IND LTD.
XX
PA      (NAKA/) NAKAMURA Y.
XX
PI      Nakamura Y, Arakawa H, Tanaka H;
XX
WPI; 2001-112446/12.
XX
FT      Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
PT      diagnosis, treatment and prevention of cancer.
XX
PS      Example 2; Page 91; 102pp; Japanese.
XX
CC      The present invention describes a human ribonucleotide reductase
CC      designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of
CC      the DNA repair mechanism and its activity is induced by p53. It can be
CC      used for the treatment, prevention and diagnosis of a wide range of
CC      cancers. The present sequence represents a human ribonucleotide reductase
CC      related sequence which is used in an example from the present invention
XX
SQ      Sequence 1081 BP; 322 A; 199 C; 252 G; 308 T; 0 U; 0 Other;
Query Match          99.8%; Score 1051.4; DB 4; Length 1081;
Best Local Similarity 99.9%; Pred. No. 2.9e-250;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy      1  ATGGCGACCGCGAAGCGCGGAGCGCGCGGCTGGATCAGATGAGATCATCTTCA 60
Db      20  ATGGCGACCGCGAAGCGCGGAGCGCGCGGCTGGATCAGATGAGATCATCTTCA 79
Qy      61  GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCCTTAAGAAAGATTCTGCG 120
Db      80  GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCCTTAAGAAAGATTCTGCG 139
Qy      121 CGGTTTCTCATCTTTCCTCAATCCAGTACCTGATATTTGGAAATGTATAAAGCGCACAG 180
Db      140 CGGTTTGTATCTTTCCTCAATCCAGTACCTGATATTTGGAAATGTATAAAGCGCACAG 199
Qy      181 GCTTCTCTCTGGACAGCAGAGAGGTGCTGATTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db      200 GCTTCTCTCTGGACAGCAGAGAGGTGCTGATTTATCAAGAGATCTCCCTCACTGGAACAAG 259
Qy      241 CTTAAAGCAGATGAGAGTACTTCAATCTCTCAGATCTTATAGCCTTTTTCAGGCGAGTGAT 300
Db      260 CTTAAAGCAGATGAGAGTACTTCAATCTCTCAGATCTTATAGCCTTTTTCAGGCGAGTGAT 319
Qy      301  GGAATTGTAAATGAAAAATTTGGTGGAGCGCTTTAGTCAAGAGGTCAGGTTCCAGAGGCT 360
Db      320  GGAATTGTAAATGAAAAATTTGGTGGAGCGCTTTAGTCAAGAGGTCAGGTTCCAGAGGCT 379
Qy      361  CGCTGTTTCTATGCGCTTCAAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTTG 420
Db      380  CGCTGTTTCTATGCGCTTCAAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTTG 439
Qy      421  CTGATAGACACTTATCATCAGAGATCCCAAGAAAAAGGGAATTTTATTAATCAATTTGAA 480
Db      440  CTGATAGACACTTATCATCAGAGATCCCAAGAAAAAGGGAATTTTATTAATCAATTTGAA 499
Qy      481  ACCATGCGCTATGTTAAGAAAAAAGCAGATTGGGCTTTGCGATGGATGATAGCAGATAGAAA 540
Db      500  ACCATGCGCTATGTTAAGAAAAAAGCAGATTGGGCTTTGCGATGGATGATAGCAGATAGAAA 559
Qy      541  TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
Db      560  TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 619
Qy      601  TCTTTTGTCTGTATATTTCTGGCTAAAGAAAGAGAGTCTTTATGCCAGGACTCACTTTTTC 660
Db      620  TCTTTTGTCTGTATATTTCTGGCTAAAGAAAGAGAGTCTTTATGCCAGGACTCACTTTTTC 679
Qy      661  AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGTGATGTTTCCAA 720
Db      680  AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGTGATGTTTCCAA 739
Qy      721  TACTTAGTAAATAAGCCTTTCAGAAAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA 780
Db      740  TACTTAGTAAATAAGCCTTTCAGAAAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA 799
Qy      781  ATTGAGCAGGAGTGTAAACAGAAAGCCTTGGCAGTTGGCCTCATTTGGAATGAATTGCATT 840
Db      800  ATTGAGCAGGAGTGTAAACAGAAAGCCTTGGCAGTTGGCCTCATTTGGAATGAATTGCATT 859
Qy      841  TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTAATTGTGGAACTTGGATTTCTCA 900
Db      860  TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTAATTGTGGAACTTGGATTTCTCA 919
Qy      901  AAGGTTTTTTCAGGCGAAGAAATCCTTTTGAATTTTATGAAACATTTCTTTAGAGGAAAA 960
Db      920  AAGGTTTTTTCAGGCGAAGAAATCCTTTTGAATTTTATGAAACATTTCTTTAGAGGAAAA 979
Qy      961  ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGCTTTTGCAGTTATATGCGAGAAC 1020
Db      980  ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGCTTTTGCAGTTATATGCGAGAAC 1039
Qy      1021 ACAGATAACGTCCTTCCACCTTGGATGCGAGATTTT 1053
Db      1040 ACAGATAACGTCCTTCCACCTTGGATGCGAGATTTT 1072
RESULT 7
AAH14924
ID      AAH14924 standard; cDNA; 1601 BP.
XX
AC      AAH14924;
XX
DT      26-JUN-2001 (first entry)
XX
DN      Human cDNA sequence SEQ ID NO:12810.
XX
KW      Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS      Homo sapiens.
XX
PN      EP1074617-A2.
XX
PD      07-FEB-2001.
XX
PF      28-JUL-2000; 2000EP-00116126.
XX

```

PR 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 XX WPI: 2001-318749/34.
 XX
 XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
 XX
 XX Claim 8; SEQ ID NO 12810; 2537pp + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention
 CC
 XX Sequence 1601 BP; 484 A; 290 C; 348 G; 479 T; 0 U; 0 Other;
 SQ

Query Match 99.7%; Score 1049.8; DB 4; Length 1601;
 Best Local Similarity 99.8%; Pred. No. 1e-289;
 Matches 1051; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCGACCGGAAAGCCGCGAGCGCGCGGCTGGATCAGATGAGATCATCTTCA 60
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 40 ATGGCGACCGGAAAGCCGCGAGCGCGCGGCTGGATCAGATGAGATCATCTTCA 99
 QY 61 GACACCAACCAAGTGAATTAAGTCAATGAAGGCCACTCTTAAGAAAGAGTTCTCGC 120
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 100 GACACCAACCAAGTGAATTAAGTCAATGAAGGCCACTCTTAAGAAAGAGTTCTCGC 159
 QY 121 CGGTTTGTCTATCTTCCAAATCCAGTACCTTGATATTTGGAATAATGATAAAGCCACAG 180
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 160 CGGTTTGTCTATCTTCCAAATCCAGTACCTTGATATTTGGAATAATGATAAAGCCACAG 219
 QY 181 GCTTCCTTCTGGACAGCAGAGAGTGCATTATCAAGGATCTCCCTCATCTGGAACAAG 240
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 220 GCTTCCTTCTGGACAGCAGAGAGTGCATTATCAAGGATCTCCCTCATCTGGAACAAG 279
 QY 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCATCTTCTTCTTCTTCTTCTTCTTCTTCT 300
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 280 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCATCTTCTTCTTCTTCTTCTTCTTCTTCT 339
 QY 301 GGAATTGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGGTTCCAGAGGCT 360
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 340 GGAATTGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGGTTCCAGAGGCT 399

QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 420
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 400 CGCTGCTTCTATGGCTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 459
 QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 480
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 460 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 519
 QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATGGGCTTTCGATGAGATGATGATGATGATGATG 540
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 520 ACCATGCCCTATGTTAAGAAAAGCAGATGGGCTTTCGATGAGATGATGATGATGATGATG 579
 QY 541 TCTACTTTTGGGGAAGAGTGTGGCTTTTGTCTGTAGAGAGGATTTTCTTCTCAGGA 600
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 580 TCTACTTTTGGGGAAGAGTGTGGCTTTTGTCTGTAGAGAGGATTTTCTTCTCAGGA 639
 QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGTCTTATGCGAGGACTCATCTTTTCTC 660
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 640 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGTCTTATGCGAGGACTCATCTTTTCTC 699
 QY 661 AATGAATCTCATCAGCAGAGATGAGGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 720
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 700 AATGAATCTCATCAGCAGAGATGAGGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 759
 QY 721 TACTTAGTAAATAAGCCTTTCAGAAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 780
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 760 TACTTAGTAAATAAGCCTTTCAGAAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 819
 QY 781 ATTGAGCAGGAGTTTTAAACAGAACCTTGCAGTTGGCCTCATTTGGAATGAATTGCATT 840
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 820 ATTGAGCAGGAGTTTTAAACAGAACCTTGCAGTTGGCCTCATTTGGAATGAATTGCATT 879
 QY 841 TTGATGAACAGTACATTTGATGTTGTAGCTGACAGATTTACTTGTGGAATCTGATCTCA 900
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 880 TTGATGAACAGTACATTTGATGTTGTAGCTGACAGATTTACTTGTGGAATCTGATCTCA 939
 QY 901 AAGGTTTTTTCAGGAGAAAATCTTTTGAATTTTATGGAATAACATTTCTTTAGAGGAAA 960
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 940 AAGGTTTTTTCAGGAGAAAATCTTTTGAATTTTATGGAATAACATTTCTTTAGAGGAAA 999
 QY 961 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGCAGAAAC 1020
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 1000 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGCAGAAAC 1059
 QY 1021 ACAGATAACGTTCTTCACTTGCAGTGCAGATTTT 1053
 DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 1060 ACAGATAACGTTCTTCACTTGCAGTGCAGATTTT 1092

RESULT 8
 ADQ57092
 ID ADQ57092 standard; DNA; 706 BP.
 AC ADQ57092;
 XX
 XX
 XX
 DT 21-OCT-2004 (first entry)
 XX
 DE Novel canine microarray-related DNA sequence SeqID8394.
 XX
 KW canine microarray; drug screening; toxicity assay;
 KW environmental pollutant; cellular response; gene expression profile;
 KW toxic response; liver necrosis; fatty liver disease;
 KW protein adduct formation; hepatitis; dog; da.
 XX
 OS Canis familiaris.
 XX
 PN WO2004063324-A2.
 XX
 PD 29-JUL-2004.
 XX
 XX 05-MAY-2003; 2003WO-US013853.
 XX
 XX 03-MAY-2002; 2002US-0377240P.

[illegible]

CC anti-diabetic; antiasthmatic; antirheumatic; antiarthritic;
 CC anti-inflammatory; antithyroid; antiallergic; antibacterial; antiviral;
 CC dermatologic; neuroprotective; cardiac; thrombolytic; coagulant;
 CC neoplastic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention
 XX
 SQ Sequence 1989 BP; 525 A; 439 C; 476 G; 544 T; 0 U; 5 Other;

Query Match 54.9%; Score 578.4; DB 3; Length 1989;
 Best Local Similarity 75.0%; Pred. No. 1.1e-154;
 Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCACTCTTAAGAAAGAGTTCTCCGCGTTTGTCTATCTTTCCATCCAGTACC 148
 DB 280 AGGATGAGCGCTGCTGAGAGAAAACCCCGCGCTTGTCTATCTTCCCATCGAGTACC 339
 QY 149 CTGATATTGGAATATGATAACAGGACAGGCTTCTTCCGACAGCAGAGAGTGC 208
 DB 340 ATGATATCTGACAGATGATTAAGAGGACAGGCTTCTTGTGGACCGCCGAGAGGTGG 399
 QY 209 ACTTATCAAGAGATCTCCCTCACTGGAACAGCTTTAAAGCAGATGAGAAGTACTTCTATCT 268
 DB 400 ACCTCTCAGAGACATTCAGCAGCTGGGAATCCCTGGAACCCGAGGAGAGATTTTATAT 459
 QY 269 CTCACATCTTAGCCCTTTTTCGACGACGATGAAATTTGAAATTTGGTGAGC 328
 DB 460 CCCATGTTCTGGCTTTCTTTCGACGACGATGATGATGATGATGATGATGATGATGATGATGAT 519
 QY 329 GCTTTAGTTCAGGAGGTGAGGTTCCAGAGGCTGCTGTTTCTATGCTTTCAAAATCTCA 388
 DB 520 GATTTAGCCAAAGATTCAGATTAACAGAACCCGCTGTTCTATGCTGCTTCAAAATGCA 579
 QY 389 TCGAGATGTTCTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 448
 DB 580 TGGAAAACATACATTTGAAATGATATAGTCTTCTTATGACACTTACATAAAGATCCCA 639
 QY 449 AGAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCGCTATGTTAAGAAAAAGCAG 508
 DB 640 AAGAAAGGGAATTTCTTCTCAATGCCATTTGAAACGATGCTTGTCTCAGAGAGAGCAG 699
 QY 509 ATTGGGCTTGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 568
 DB 700 ACTGGGCTTGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 759
 QY 569 TTGCTGCTGATAGAGGATTTTCTTCTCAGAGATTTTCTGCTGCTATATTTCTGCTCAAGA 628
 DB 760 TTGCTGCTGATGAGGATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 819
 QY 629 AGAGAGTCTTATTCACGAGTCTCACTTTTCCATGAACTCATCAGCAGAGATGAGGAC 688
 DB 820 AACGAGACTGATGCTTGGCTTCACTTTCTAATGAACTTATGACAGATGAGGCTT 879
 QY 689 TTCACTGCTGATTTGCTTGGCTGATGTTTCAATATCTTAGTAAATTAAGCTTTCAAGAGAA 748
 DB 880 TACACTGCTGATTTGCTTGGCTGATGTTTCAACACCTGTTACACAAACCATCGAGGAGA 939
 QY 749 GGGTCAGGAGATCAATTTGATGCTGCTCAAAATTTGAGCAGGAGTCTTTTACAGAGACCT 808
 DB 940 GAGTAAGAGAAATTAATTTATCAATGCTGTTGGATGAGAACAGAGGTTCTCTCACTGAGGCT 999
 QY 809 TGCCAGTTGGCTCTCATTTGGAATGAATTTTGTGATGAAACAGTACATTTGATGTTTGTAG 868

DB 1000 TGCTGTGAAGCTCATTTGGGATGAATTCGACTCTAATGAAGCAATACATTTGATTTGTG 1059
 QY 869 CTGACAGATTACTTGTGGAACCTTGATTTCTCAAGGTTTTTCAGGCAGAAAAATCCTTTTG 928
 DB 1060 CAGACAGACTTATGCTGGAACCTTGGTTTATGCAAGGTTTTTCAGATGAGAACCCATTTG 1119
 QY 929 ATTTTATGAAAAACATTTCTTTTAGAAGGAAAAACAAATTTCTTTGAGAAAAACGAGTTTCAG 988
 DB 1120 ACTTTATGAGAAATTTTCTTCTGGAAGGAAGACTTAACCTTTCTTGAAGAGAGTAGGCG 1179
 QY 989 AGTATCAGCGTTTTCAGTTATGCGAGAAACCAAGATTAACCTTTCTTACCTTTGAGATGAG 1048
 DB 1180 AGTATCAGAGGATGAGGAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1239
 QY 1049 ATTT 1052
 DB 1240 ACTT 1243

RESULT 10
 ADK70302

ID ADK70302 standard; cDNA; 2215 BP.

XX AC ADK70302;

XX DT 06-MAY-2004 (first entry)

XX DE Respiratory disease differentially expressed cDNA #38.

XX KW ds; gene; cytostatic; respiratory; antiasthmatic; gene therapy;
 XX KW differential gene expression; respiratory disorder; lung cancer;
 XX KW chronic obstructive pulmonary disease; emphysema; asthma.

XX OS Homo sapiens.

XX FN WO2003101283-A2.

XX PD 11-DEC-2003.

XX PF 02-JUN-2003; 2003WO-US017409.

XX PR 04-JUN-2002; 2002US-0386005P.

XX XX (INCY-) INCYTE CORP.

XX XX Rickert PK, Krasnow R;

XX XX WPI; 2004-042945/04.

XX PS Claim 1; SEQ ID NO 38; 343pp; English.
 XX CC The invention relates to cDNA sequences that are differentially expressed
 CC in respiratory disorders or their complements or encoded proteins. The
 CC cDNAs and proteins are useful for diagnosing, treating or monitoring
 CC treatment of a subject with a respiratory disease including lung cancer,
 CC chronic obstructive pulmonary diseases, emphysema or asthma. The protein
 CC is also useful for screening molecules or compounds to identify at least
 CC one ligand which specifically binds the protein. It is also useful for
 CC preparing and purifying a polyclonal or monoclonal antibody. This
 CC sequence corresponds to a cDNA of the invention.

XX SQ Sequence 2215 BP; 593 A; 476 C; 523 G; 623 T; 0 U; 0 Other;

XX CC Query Match 54.9%; Score 578.4; DB 12; Length 2215;
 XX CC Best Local Similarity 75.0%; Pred. No. 1.2e-154;
 XX CC Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

Db 464 CCCAGTCTCGCTTCTTTGAGCAAGCGATGGCATAGTAATGAAATCTGGTGGAGC 523
Qy 329 GCTTTAGTCAGAGGTGAGGTTCAGAGGCTCGCTTTCTATGGCTTTCAAAATCTCA 388
Db 524 GATTAGCAGAAGTTCAGATTACAGAACCCGCTGTTCTATGGCTTCCAAATTCGCA 583
Qy 389 TCGAAGATGTTCTACTCAGAGATGTACAGTTTGTCTGATAGACACCTTACATCAGAGATCCCA 448
Db 584 TGGAAAACATACATCTCTGAATGTATAGTCTTCTTTATTTGACACTTACATAAAGATCCCA 643
Qy 449 AGAAAGGGAATTTTATTTAATGCAATGAAACCATGCGCTATGTTTAAGAAAAAGCAG 508
Db 644 AAGAAAGGGAATTTCTCTTCAATGCCATTGAACGATGCTTGTGTCAAGAAAGAGGCAG 703
Qy 509 ATTGGGCTTCGCGATGATAGACAGATGAAATCTACTTTTGGGGAAGAGTGGTGGCT 568
Db 704 ACTGGGCTTCGCGATGATGATGAGGCAAGAGGCTTACCTATGTTGGAACGTGTGTAGGCT 763
Qy 569 TTGCTGCTGTAGAAGGAGTCTTCTCTCAGGATCTTTTGGCTGCTATATTTCTGGCTAAAGA 628
Db 764 TTGCTGAGTGAAGGCAATTTCTTTCCGGTCTTTTGGCTGCTATATTTCTGGCTCAAGA 823
Qy 629 AGAGAGTCTTATGCGCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGCAG 688
Db 824 AACGAGGACTGATGCTGGCCTCACATTTTCTAATGAACTTATTAGCAGAGATGAGGCTT 883
Qy 689 TTCACTGTGATTTGCTGCTGATGTTCCAAATGAACTTATTAGTAAATGAAGCTTCAGAGAAA 748
Db 884 TACACTGTGATTTTGGCTGCTGATGTTCAAAACACCTGGTACACAAACCATCGGAGGAGA 943
Qy 749 GGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTTGAGCAGGAGTTTAAACAGAGCCT 808
Db 944 GAGTAAGGAATTAATATCAATGCTGTTCCGATAGAACAGAGTCTCTACTGAGGCTT 1003
Qy 809 TGCCAGTTGGCCTCATTTGGAATGAATTCATTTTGAAGAAACAGTACATTCAGTTTGTAG 868
Db 1004 TGCTGTGAGACTCATTTGGGATGAATTCGACTTAATGAACCAATACATTCAGTTTGTG 1063
Qy 869 CTGACAGATTAATGTTGGAATGATGATCTCAAGGTTTTCAGGCGAGAAATCTTTTG 928
Db 1064 CAGACAGACTTATGCTGGAATGCGTTTGAAGAGTTTTCAGAGTGAAGAACCAATTTG 1123
Qy 929 ATTTTATGGAACATTTCTTAGAGGAAACCAATTTCTTTGAGAAACGCTTTCAG 988
Db 1124 ACTTATGGAATTAATTTCTAGGAAAGAACTAACTCTTTTGAAGAGAGTAGGCG 1183
Qy 989 AGTATCAGCGTTTTCAGTTATGTCAGAGAAACCAAGATAAGCTTCTACCTTGCATGCAG 1048
Db 1184 AGTATCAGAGGATGGGATGATGTCAGTCCACAGAGAAATCTTTTACCTTGGATGCTG 1243
Qy 1049 ATTT 1052
Db 1244 ACTT 1247

RESULT 12

AAS44917/C
ID AAS44917 standard; DNA; 2482 BP.

XX AC AAS44917;

XX DT 18-DEC-2001 (first entry)

XX Human contig polynucleotide sequence #170.

XX Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
XX mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
XX cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
XX nervous system disorder; inflammatory disorder; cell differentiation; ds;
XX angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
XX genetic disorder; bone regeneration; tendon; ligament; tissue repair;
XX cytosstatic; antirheumatic; antiarthritic; vulnery; antiinflammatory;

KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
KW neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;
KW immunostimulant; analgesic; gene therapy.

OS Homo sapiens.
OS Synthetic.

XX WO200164834-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004926.

XX 28-FEB-2000; 2000US-00515126.

PR 18-MAY-2000; 2000US-00577409.

PR 17-JUN-2000; 2000US-00597707.

PR 14-JUL-2000; 2000US-00616807.

PR 19-SEP-2000; 2000US-00664641.

XX (HYSE-) HYSEQ INC.

Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;
Drmanac R;

WPI: 2001-589862/66.

P-PSDB; AAU28017.

Novel polypeptides and nucleic acids obtained from cDNA libraries

prepared from various human tissues, for diagnosis, treatment of cancer,

neurological, inflammatory disorders and for use in arrays for detection.

Claim 1; SEQ ID NO 514; 153pp; English.

Sequences AAS44576-AAS44919 represent full-length polynucleotides and

contig polynucleotides encoding polypeptides of the invention. The DNA

and protein sequences are useful for the treatment, diagnosis and

prevention of various types of disorder in a mammalian subject such as a

human, dog, monkey, mouse, hamster or rat. The disorders include cancers

such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such

as multiple sclerosis, connective tissue disease, rheumatoid arthritis,

diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system

disorders such as Parkinson's disease, Alzheimer's disease, Huntington's

chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and

Wernicke disease, inflammatory disorders such as nephritis, Crohn's

disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory

bowel disease. The sequences exhibit activity relating to angiogenesis,

cell proliferation, cell differentiation, stem cell growth factor,

activin or inhibin. Therefore, they can be used to manipulate stem cells

in culture to give rise to neuroepithelial cells that can be used to

augment or replace cells damaged by illness, accidental damage or genetic

disorders. The sequences may also be used for regeneration of bone,

cartilage, tendons and ligaments and in tissue repair and burn healing.

Note: Some sequences for this patent did not form part of the printed

specification, but were obtained in electronic format directly from WIPO

at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 2482 BP; 707 A; 598 C; 522 G; 655 T; 0 U; 0 Other;

Query Match 54.9%; Score 578.4; DB 4; Length 2482;

Best Local Similarity 75.0%; Pred. No. 1.2e-154;

Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

Qy 89 ATGAAGAGCCACTCTCTAAGAAGAGTCTCGCGGTTTGTCTATCTTTTCCAAATCCAGTACC 148

Db 2086 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTCTATCTTTCCCATCGAGTACC 2027

Qy 149 CTGATATTTGAAATATGATATAACAGGCACAGGCTTCTCTTGGACAGCAGAGAGTGC 208

Db 2026 ATGATATCTGCAGATGATATAAGAGGCGAGGCTTCTCTTTGGACCGCGAGAGTGC 1967

Qy 209 ACTTATCAAGAGATCTCCCTCAGTCGAAACAGCTTAAGACAGATGAGAGTACTTCATCT 268

Db 1966 ACCTCTCCAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 1907
 Qy 269 CTCACTCTTAGCCTTTTTCAGCCAGTATGTAATTTGTAATTTGTTGGAGC 328
 Db 1906 CCCATGTTCTGGCTTTCTTTCAGCAGGAGGATGATGTAATTTGTTGGAGC 1847
 Qy 329 GCTTTAGTCAGGAGGTCAGGTTCCAGAGGCTCGCTGTTCTATGGCTTTCAAAATTTCTCA 388
 Db 1846 GATTTAGCCAGAGATTCAGATTACAGAACCCGCTGTTCTATGGCTTCCAAATTTGCCA 1787
 Qy 389 TCGAGATGTTCACTCAGAGATGATAGATTTGCTGATAGACACTTATCAGAGATCCCA 448
 Db 1786 TGGAAAACATACATTTGAAATGATATGCTTCTTATTTGACACTTACATAAAGATCCCA 1727
 Qy 449 AGAAAAGGGAATTTTATTTAATCAATGGAACCATGCGCCCTATGTTAAGAAAAGAGCAG 508
 Db 1726 AAGAAAGGAATTTCTCTTCAATGCCATTTGAACGATGCTTGTCTAAGAAAGGAGCAG 1667
 Qy 509 ATTGGGCTTTGCGATGATAGCAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCT 568
 Db 1666 ACTGGGCTTTGGCTGGATGGGACAAAGAGGCTACTATGGTGAACGTTGTAGGCT 1607
 Qy 569 TTGCTGCTGATAGAGGATTTTCTTCTCAGAGCTTTTTCGCTATATTTCTGCTTAAGA 628
 Db 1606 TTGCTGAGTGAAGGCAATTTTCTTTCGGCTTCTTTCGCTCGATATTTCTGGCTCAAGA 1547
 Qy 629 AGAGAGTCTATGCCAGGACTCACTTTTCCATGAACTCATCAGCAGAGATCAAGGAC 688
 Db 1546 AACGAGGACTGATGCGCTGCTCACTTTCTAATGAACCTTATTAGCAGAGATGAGGTT 1487
 Qy 689 TTCACTGTGACTTTGCTTGGCTGATGTTTCAATATCTTAGTAAATAAGCCTTCAAGAGAA 748
 Db 1486 TACACTGTGATTTTGTCTGCTGATGTTCAACACCTGTTACACACCATCGAGGAGA 1427
 Qy 749 GGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTTGAGCAGAGATTTTAAACAGAGCCT 808
 Db 1426 GAGTAAGAGAAATAATATCAATGCTGTTCCGATAGAAACAGGAGTTCCTACTGAGGCT 1367
 Qy 809 TGCAGTTGGCCTCATTTGGATGATGATTTTGTATGAACACATCATTTGAGTTTGTAG 868
 Db 1366 TGCCTGTGAAGCTCATTTGGGATGAATTTGCACTCTTAATGAAGCAATACATTTGAGTTGTGG 1307
 Qy 869 CTGACAGATTAATTTGGAATCTGATGTTTCAAAAGTTTTCAGGCGAGAAATCCTTTTG 928
 Db 1306 CAGACAGACTTATGCTGGAATCTGGGTTTACCAAGTTTTCAGAGTAGAGACCATTTG 1247
 Qy 929 ATTTATGGAATAACATTTCTTTAGAGAGAAACAAATTTCTTTGAGAAACGAGTTTTCAG 988
 Db 1246 ACTTTATGGAGATATTTCACTGGAAGGAAAGACTAACTTCTTTGAGAGAGAGTAGGCG 1187
 Qy 989 AGTATCAGCGTTTTCAGTTATGCGAGAAACCAACAGATAAGCTTTCACCTTGGATGCG 1048
 Db 1186 AGTATCAGAGATGGGAGTGTATGTCAGTCCACAGAGAAATTTCTTTTACCTTGGATGCTG 1127
 Qy 1049 ATTT 1052
 Db 1126 ACTT 1123

RESULT 13
 ABL65414
 ID ABL65414 standard; DNA; 2500 BP.
 XX AC ABL65414;
 XX XX
 DT 15-MAY-2002 (first entry)
 XX XX
 DE Lung cancer related gene sequence SEQ ID NO:3751.
 XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; ds.

XX Homo sapiens.
 OS WO200194629-A2.
 PN 13-DEC-2001.
 FD 30-MAY-2001; 2001WO-US010838.
 XX 05-JUN-2000; 2000US-0209473P.
 XX 05-JUN-2000; 2000US-0209531P.
 PR 18-SEP-2000; 2000US-0233133P.
 PR 18-SEP-2000; 2000US-0233617P.
 PR 20-SEP-2000; 2000US-0234009P.
 PR 20-SEP-2000; 2000US-0234034P.
 PR 20-SEP-2000; 2000US-0234052P.
 PR 22-SEP-2000; 2000US-0234509P.
 PR 22-SEP-2000; 2000US-0234567P.
 PR 25-SEP-2000; 2000US-0234924P.
 PR 25-SEP-2000; 2000US-0235077P.
 PR 25-SEP-2000; 2000US-0235082P.
 PR 25-SEP-2000; 2000US-0235134P.
 PR 25-SEP-2000; 2000US-0235280P.
 PR 26-SEP-2000; 2000US-0235637P.
 PR 26-SEP-2000; 2000US-0235638P.
 PR 27-SEP-2000; 2000US-0235711P.
 PR 27-SEP-2000; 2000US-0235720P.
 PR 27-SEP-2000; 2000US-0235840P.
 PR 28-SEP-2000; 2000US-0236028P.
 PR 28-SEP-2000; 2000US-0236032P.
 PR 28-SEP-2000; 2000US-0236033P.
 PR 28-SEP-2000; 2000US-0236034P.
 PR 28-SEP-2000; 2000US-0236110P.
 PR 29-SEP-2000; 2000US-0236842P.
 PR 29-SEP-2000; 2000US-0236891P.
 PR 02-OCT-2000; 2000US-0237172P.
 PR 02-OCT-2000; 2000US-0237173P.
 PR 02-OCT-2000; 2000US-0237278P.
 PR 02-OCT-2000; 2000US-0237294P.
 PR 02-OCT-2000; 2000US-0237295P.
 PR 03-OCT-2000; 2000US-0237316P.
 PR 03-OCT-2000; 2000US-0237425P.
 PR 03-OCT-2000; 2000US-0237598P.
 PR 03-OCT-2000; 2000US-0237604P.
 PR 03-OCT-2000; 2000US-0237606P.
 PR 01-NOV-2000; 2000US-0244867P.
 PR 01-NOV-2000; 2000US-0245084P.
 XX (AVAL-) AVALON PHARM.
 XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX WPI; 2002-188264/24.

Screening for anti-neoplastic agent involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, and determining a change in expression of a gene of a signature gene set.

Claim 1; SEQ ID NO 3751; 44pp; English.

The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in ABL61664 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an

CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of MI, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. MI can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;

Query Match 54.9%; Score 578.4; DB 6; Length 2500;
Best Local Similarity 75.0%; Pred. No. 1.2e-154;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

Qy	89	ATGAAGCCCACTCTCAGAAAGAGTTCTCGCGGTTTGTCATCTTTCCATCCAGTACC	148
Db	397	AGGATGACCGCTGCTGAGAGAAACCCCGCGCTTTGTCACTTCCCATCGAGTACC	456
Qy	149	CTGATATTGGAAATGATATAACAGGACAGGCTTCTCTGGACAGCAGAGAGTGC	208
Db	457	ATGATATCTGCAGATGATAGAGAGCAGAGCTTCTTTGGACCCCGAGGAGTTG	516
Qy	209	ACTTATCAAGAGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTACTTCATCT	268
Db	517	ACCTCTCAAGAGACATTGAGCACTGGGAATCCCTGNAACCCGAGGAGATATTTATAT	576
Qy	269	CTCAGATCTTAGCTTTTTCAGCCAGTATGGAATTTGTAATGAAATTTGGTGAGC	328
Db	577	CCCATGTTCTGGCTTTCTTTGACGACGATGGCATAGTAAATGAAATCTGGTGAGC	636
Qy	329	GCTTTAGTCAGAGTGAGTTCAGAGGCTCGCTGTTCTTATGGCTTTCAAAATCTCA	388
Db	637	GATTTAGCCAGAGAGTTCAGATTAAGAGCCCGCTGTTCTATGGCTTCCAAATGCGCA	696
Qy	389	TCGAGATGTTTCACTCAGAGATGTACAGTTTGTGATAGACACTTTACATCAGAGATCCCA	448
Db	697	TGGAACACATACATCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA	756
Qy	449	AGAAAGGGAATTTTATTAATGCAATGAAACCATGCCCTATGTTAAGAAAAAAGCAG	508
Db	757	AAGAAAGGGAATTTCTCTCAATGCCATGGAACGATGCTTGTCTCAAGAAAGGCGAG	816
Qy	509	ATTGGGCTTGGATGATAGACATAGAAATCTACTTTTGGGGAAGAGTGTGGCCT	568
Db	817	ACTGGGCTTGGCTGATGAGGACAAAGAGGCTACCTATGGTGAACGTTTGTAGCCT	876
Qy	569	TTGCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTGTGCTATATTCTGGCTAAAGA	628
Db	877	TTGCTGAGTGAAGGCAATTTCTTTTCCGGTCTTTTGGTTCGATATTTCTGGCTCAGA	936
Qy	629	AGAGAGTCTTATGCCAGGACTCACTTTTTCATGAACTCATCAGCAGAGATGAAGAC	688
Db	937	AACGAGGACTGATGCCCTGCCCTACATTTTCTAATGAATTTATTAGCAGATGAGGTT	996
Qy	689	TTCACTGTGATTTGCTTGGCTGATGTTCCATATCTAGTAAATAGCCTTCAGAGAAA	748
Db	997	TACACTGTGATTTTGGCTTGGCTGATGTTTCAACACCTGGTACACAAACCATCGGAGGA	1056
Qy	749	GGGTGAGGAGATCATTTGTTGATGCTGTCATAAATTTGAGCAGGAGTTTAAACAGCCT	808
Db	1057	GAGTAAAGAAATTAATTAATCATGCTGTTCGATAGAACAGGAGTTCCTACTGAGGCT	1116
Qy	809	TGCCAGTTGGCTCATTTGGAATGAATTCATTTTGTATGAACAGTACATTCAGTTTGTAG	868
Db	1117	TGCTGTGAAGCTCATTTGGGATGAATTTGCACTCTAATGAAGCAATACATTGAGTTTGTG	1176
Qy	869	CTGACAGATTAATTTGTTGGAATCTTCAAGAGTTTTCAGGCGAGAAATCTTTTG	928
Db	1177	CAGACAGACTTATGCTGGAATCTGGGTTTATGCAAGGTTTTCAGAGTAGAGAACCAATTG	1236
Qy	929	ATTTTATGAAAAACATTTCTTTAGAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAG	988

Db	1237	ACTTTATGAGAAATATTTCACTGGAGGAAAGACTAACTTCTTTGAGAGAGATAGCG	1296
Qy	989	AGTATCAGCGTTTTCAGTTATGCGAGAAACCAACAGATAACGTTCTTACCTTCGATGCG	1048
Db	1297	AGTATCAGAGGATGGAGTATGTCAGTCCACACAGAGAAATCTTTTACCTTCGATGCG	1356
Qy	1049	ATTT 1052	
Db	1357	ACTT 1360	
RESULT 14			
ABL66517			
ID	ABL66517 standard; DNA; 2500 BP.		
XX	AC ABL66517;		
XX	XX		
DT	15-MAY-2002 (first entry)		
XX	Lung cancer related gene sequence SEQ ID NO:4854.		
DE	Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;		
KW	stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;		
KW	cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;		
KW	gene; da.		
XX	XX		
OS	Homo sapiens.		
XX	XX		
PN	WO200194629-A2.		
XX	XX		
PD	13-DEC-2001.		
XX	XX		
PF	30-MAY-2001; 2001WO-US010838.		
XX	XX		
PR	05-JUN-2000; 2000US-0209473P.		
PR	05-JUN-2000; 2000US-0209531P.		
PR	18-SEP-2000; 2000US-0233133P.		
PR	18-SEP-2000; 2000US-0233617P.		
PR	20-SEP-2000; 2000US-0234009P.		
PR	20-SEP-2000; 2000US-0234034P.		
PR	22-SEP-2000; 2000US-0234509P.		
PR	22-SEP-2000; 2000US-0234567P.		
PR	25-SEP-2000; 2000US-0234923P.		
PR	25-SEP-2000; 2000US-0234924P.		
PR	25-SEP-2000; 2000US-0235077P.		
PR	25-SEP-2000; 2000US-0235082P.		
PR	25-SEP-2000; 2000US-0235134P.		
PR	25-SEP-2000; 2000US-0235280P.		
PR	26-SEP-2000; 2000US-0235637P.		
PR	27-SEP-2000; 2000US-0235638P.		
PR	27-SEP-2000; 2000US-0235711P.		
PR	27-SEP-2000; 2000US-0235720P.		
PR	27-SEP-2000; 2000US-0235840P.		
PR	28-SEP-2000; 2000US-0235863P.		
PR	28-SEP-2000; 2000US-0236028P.		
PR	28-SEP-2000; 2000US-0236032P.		
PR	28-SEP-2000; 2000US-0236033P.		
PR	28-SEP-2000; 2000US-0236034P.		
PR	28-SEP-2000; 2000US-0236109P.		
PR	28-SEP-2000; 2000US-0236111P.		
PR	29-SEP-2000; 2000US-0236842P.		
PR	29-SEP-2000; 2000US-0236891P.		
PR	02-OCT-2000; 2000US-0237172P.		
PR	02-OCT-2000; 2000US-0237278P.		
PR	02-OCT-2000; 2000US-0237294P.		
PR	02-OCT-2000; 2000US-0237295P.		
PR	02-OCT-2000; 2000US-0237316P.		
PR	03-OCT-2000; 2000US-0237425P.		
PR	03-OCT-2000; 2000US-0237598P.		
PR	03-OCT-2000; 2000US-0237604P.		

PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
XX
XX (AVAL-) AVALON PHARM.
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
PI
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
XX
PS Claim 1; SEQ ID NO 4854; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
XX Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;

Query Match 54.9%; Score 578.4; DB 6; Length 2500;
Best Local Similarity 75.0%; Pred. No. 1.2e-154;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCACTCTAAGAAAGAGTTCTCCCGGTTGTCTCTTCCAAATCCAGTACC 148
DB |||||
QY 149 CTGATATTGGAAATGTATAAACAGGCACAGGCTTCTTCTGGACAGCAGAGGTG 208
DB |||||
QY 209 ACTTATCAAGGATCTCCCTCACTGGACACAGCTTAAAGCAGATGAGAGTACTTCACT 268
DB |||||
QY 517 ACCTCTCAAGGACATTCAGCACTGGGAAATCCCTGNAACCCGAGAGAGATATTTATAT 576
QY 269 CTCACTATTAGCTTTTGGACGAGTGAATTTGAAATGAAATTTGGTGAGC 328
DB |||||
QY 577 CCATGTTCTGGCTTCTTGGCAGCAAGCGATGCGATGATGAAATCTTGGTGAGC 636
QY 329 GCTTTAGTCAGGAGGTGAGGTTCCAGAGGCTCGCTGTGTTCTATGGCTTTCAAAATCTCA 388
DB |||||
QY 637 GATTTAGCCAAAGATTTCAGATTACAGAACCCGCTGTTCTATGGCTTCCAAATTGCA 696
QY 389 TCGAATGTTTCACTCAGAGATGTACAGTTTGTGATGACACTTACATCAGATGCCA 448
DB |||||
QY 697 TGGAAACATACATCTCTGAAATGTATAGTCTTCTTATTGACACTTACATAAAGATCCCA 756
QY 449 AGAAAGGGAAATTTTATTAATGCAATGAAACCATGCAATGTTTAAAGAAAGACAG 508
DB |||||
QY 757 AAGAAAGGGAAATTTCTCTTCAATGCCATTTGAAACGATGCGCTTGTGTCAAGAAAGGCGAG 816
QY 509 ATTGGGCTTGGCATGATGATGACATAGAAATCTACTTTTGGGAAAGAGTGGTGCGCT 568
DB |||||
QY 817 ACTGGGCTTGGCTGATTTGGGACAAAGAGGCTACTATGTTGTAAGCTGTTGAGCT 876
DB |||||

QY 569 TTGCTGCTGTAGAAGAGTTTCTTCTCAGGATCTTTTGTGCTATATTTCTGCTAAAGA 628
DB |||||
QY 877 TTGCTGAGTGAAGGCATTTTCTTTCCGGTCTTTTCCGTCGATATTTCTGGCTCAAGA 936
DB |||||
QY 629 AGAGAGGTCTTATGCCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGGAC 688
DB |||||
QY 937 AACGAGGACTGATGCTGGCCTCACTTTTCTAATGAACTTATTAGCAGAGATGAGGGTT 996
DB |||||
QY 689 TTCACTGTGACTTTGCTTGCCTGATGTTCCAAATGAACTTATAGTAAATGAGCCTTCAGAGAAA 748
DB |||||
QY 997 TACACTGTGATTTTGTCTTGCCTGATGTTTCAACACCTGGTACACAAACCATCGGAGGAGA 1056
QY 749 GGGTCAGGAGCATCACTTGTGATGCTGTCAAAATTTGAGCAGGAGTTTAAACAGAGGCT 808
DB |||||
QY 1057 GAGTAAGAGAATTAATATCAATGCTGTTCGGATAGAACAGAGGTTCCTCCTGAGGCT 1116
QY 809 TGCCAGTTGGCCTCAATTGGAATGAATTCATTTTGAAGGAAACAAATTTCTTTCAGAGAAACGAGTTTTCAG 868
DB |||||
QY 1117 TGCCTGTGAAGCTCAATTGGGATGAATTCGACTCTAATGAAGCAATACATTTGAGTTTGTGG 1176
QY 869 CTGACAGATTAATTTGGAAGTCTTCAAGGTTTTCAGGAGGAGGAAATCCTTTTG 928
DB |||||
QY 1177 CAGACAGACTTATGCTGGAAGTGGGTTTTCAGAGGTTTTCAGAGTAGAGAAACCCATTTG 1236
QY 929 ATTTTATGGAACCAATTTCTTTTGAAGGAAACAAATTTCTTTCAGAGAAACGAGTTTTCAG 988
DB |||||
QY 1237 ACTTTATGGAATATTTTCACTGGAAGGAAAGACTTAATCTTTGAGAGAGAGTAGGCG 1296
QY 989 AGTATCAGCGTTTTCAGTATTTGAGGAGAAACCCACAGATAACGTTCTTTCACCTTGGATGAC 1048
DB |||||
QY 1297 AGTATCAGAGGATGGAGTGTCAAGTCCAAACAGAGAAATTTCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
DB |||||
QY 1357 ACTT 1360

RESULT 15
ABL65859
ID ABL65859 standard; DNA; 2500 BP.
XX
AC ABL65859;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:4196.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
XX WO200194629-A2.
XX
XX 13-DEC-2001.
XX
XX 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 26, 2005, 00:00:10 ; Search time 160 Seconds
(without alignments)
10768.746 Million cell updates/sec

Title: US-10-698-228-2
Perfect score: 1053
Sequence: 1 atgggagccggaaggcc.....tcacctggatgcagatttt 1053

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgm2_6/ptodata/1/ina/5A COMB.seq:*
- 2: /cgm2_6/ptodata/1/ina/5B COMB.seq:*
- 3: /cgm2_6/ptodata/1/ina/6A COMB.seq:*
- 4: /cgm2_6/ptodata/1/ina/6B COMB.seq:*
- 5: /cgm2_6/ptodata/1/ina/PCUTS COMB.seq:*
- 6: /cgm2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1053	100.0	1053	4	US-10-019-733-2
2	1053	100.0	4955	4	US-10-019-733-3
3	1051.4	99.8	1053	4	US-10-019-733-12
4	1051.4	99.8	1081	4	US-10-019-733-4
5	579	55.0	2500	4	US-09-962-665-9
6	579	55.0	2500	4	US-09-962-677-9
7	579	55.0	2500	4	US-09-949-016-2025
8	578.4	54.9	2479	4	US-09-023-655-1370
9	578.4	54.9	2500	4	US-09-949-016-145
10	578.4	54.9	2500	4	US-09-949-016-145
11	441	41.9	481	3	US-08-905-223-125
12	429.8	40.8	14176	1	US-08-307-499-1
13	429.8	40.8	14176	1	US-08-307-499-14
14	429.8	40.8	14176	3	US-09-299-268-1
15	429.8	40.8	14176	3	US-09-299-268-14
16	382.2	36.4	1236	4	US-09-248-796A-3914
17	382.8	36.4	825	1	US-08-307-499-40
18	382.8	36.4	825	3	US-09-299-268-40
19	337.8	32.1	1112	1	US-08-136-743B-1
20	318.2	30.2	419	3	US-08-905-223-129
21	316	30.0	1083	4	US-09-248-796A-3915
22	260.8	24.8	656	3	US-09-328-111-791
23	195.2	18.5	449	4	US-09-270-767-14670
24	169.2	16.1	366	1	US-08-307-499-13
25	169.2	16.1	366	3	US-09-299-268-13
26	141.4	13.4	373	4	US-09-401-064-337
27	139.8	13.3	373	4	US-09-401-064-319

c	28	115.2	10.9	366	4	US-09-401-064-307	Sequence 307, App
	29	112.4	10.7	297	4	US-09-313-294A-4435	Sequence 4435, Ap
	30	93.8	8.9	11820	4	US-09-949-016-11887	Sequence 11887, A
	31	93.8	8.9	11826	4	US-09-949-016-13767	Sequence 13767, A
	32	92	8.7	601	4	US-09-949-016-20615	Sequence 20615, A
	33	92	8.7	601	4	US-09-949-016-69028	Sequence 69028, A
	34	74.4	7.1	276	4	US-09-313-294A-3437	Sequence 3437, Ap
	35	70.2	6.7	304	4	US-09-313-294A-7143	Sequence 7143, Ap
	36	68	6.5	276	4	US-09-313-294A-4638	Sequence 4638, Ap
	37	61	5.8	1053	4	US-09-902-540-2666	Sequence 2666, Ap
c	38	61	5.8	13706	4	US-09-902-540-1124	Sequence 1124, Ap
	39	60	5.7	601	4	US-09-949-016-69031	Sequence 69031, A
	40	59.8	5.7	1141	4	US-09-806-708B-22	Sequence 22, Appl
	41	54.4	5.2	288	4	US-09-313-294A-6108	Sequence 6108, Ap
	42	52.2	5.0	276	4	US-09-313-294A-3244	Sequence 3244, Ap
	43	50	4.7	1230025	4	US-09-198-452A-1	Sequence 1, Appli
	44	50	4.7	1230230	4	US-09-438-185A-1	Sequence 1, Appli
c	45	45.4	4.3	124884	4	US-09-661-596A-76	Sequence 76, Appl

ALIGNMENTS

RESULT 1

US-10-019-733-2
; Sequence 2, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-2

Query Match	100.0%	Score 1053;	DB 4;	Length 1053;
Best Local Similarity	100.0%	Pred. No. 8.7e-311;	Mismatches 0;	Indels 0;
Matches 1053;	Conservative 0;			Gaps 0;
Qy	1	ATGGCGCACCGGAAAGCCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA	60	
Db	1	ATGGCGCACCGGAAAGCCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA	60	
Qy	61	GACACCAAGAAAGTGAATAAAGTCAATGAAGCCACTCTTAAAGAGGTTCTCG	120	
Db	61	GACACCAAGAAAGTGAATAAAGTCAATGAAGCCACTCTTAAAGAGGTTCTCG	120	
Qy	121	CGGTTTGTCTTTCGATCCAGTACCTGATATTTGGAAATGATATTAACAGGCACAG	180	
Db	121	CGGTTTGTCTTTCGATCCAGTACCTGATATTTGGAAATGATATTAACAGGCACAG	180	
Qy	181	GTTTCTTCTTGACAGCAGAGAGGTCCAGTATTAACAGGATCTCCCTCAGTGAACAAG	240	
Db	181	GTTTCTTCTTGACAGCAGAGAGGTCCAGTATTAACAGGATCTCCCTCAGTGAACAAG	240	
Qy	241	CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACTTTAGCCCTTTTTCAGCCAGTGAT	300	
Db	241	CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACTTTAGCCCTTTTTCAGCCAGTGAT	300	
Qy	301	GGAAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTTCAGAGGCTT	360	

Db 301 GGAATTCTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
QY 361 CGCTGTTTCTATGCTTCAAAATCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGCTTCAAAATCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTTAATGCAATTTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTTAATGCAATTTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGGCATGGATAGCAGATAGAAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGGCATGGATAGCAGATAGAAAA 540
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCCCTTGGCTGCTAGAGAGGATTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGGAAGAGTGGTGGCCCTTGGCTGCTAGAGAGGATTTTCTTCTCAGGA 600
QY 601 TCTTTTCTGCTATTTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACATTTTTC 660
Db 601 TCTTTTCTGCTATTTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACATTTTTC 660
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGGCTGCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGGCTGCTGATGTTCCAA 720
QY 721 TACTTAGTAATAAGCCCTCAGAAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA 780
Db 721 TACTTAGTAATAAGCCCTCAGAAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA 780
QY 781 ATTGAGCAGAGTTTTAAACAGAGCCCTGCGAGTGCCTCATTTGGAATGAATTCATTT 840
Db 781 ATTGAGCAGAGTTTTAAACAGAGCCCTGCGAGTGCCTCATTTGGAATGAATTCATTT 840
QY 841 TTGATGAAACAGTACATTTGATGTTGTAGTCAGCAGATTTACTTTGGAACCTTGGATTTCTCA 900
Db 841 TTGATGAAACAGTACATTTGATGTTGTAGTCAGCAGATTTACTTTGGAACCTTGGATTTCTCA 900
QY 901 AAGTTTTTTCAGGAGAAAAATCCCTTTTGTATTTATGGAAGAAACATTTCTTTAGAGGAAAA 960
Db 901 AAGTTTTTTCAGGAGAAAAATCCCTTTTGTATTTATGGAAGAAACATTTCTTTAGAGGAAAA 960
QY 961 ACAAAATTTCTTTGAGAACAGTTTTCAGAGTATCAGAGTTCAGAGTTATGGCAGAGAAC 1020
Db 961 ACAAAATTTCTTTGAGAACAGTTTTCAGAGTATCAGAGTTCAGAGTTATGGCAGAGAAC 1020
QY 1021 ACAGATAACGTCCTTCACTTTGGATGCAGATTTT 1053
Db 1021 ACAGATAACGTCCTTCACTTTGGATGCAGATTTT 1053

RESULT 2
US-10-019-733-3
; Sequence 3, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 3
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens

; FEATURE:
; OTHER INFORMATION:
US-10-019-733-3
Query Match 100.0%; Score 1053; DB 4; Length 4955;
Best Local Similarity 100.0%; Pred. No. 2.1e-310;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGCGACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 245 ATGGGCGACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATCAAGAGCCACTCTCTAAGAAAGATTTCTCG 120
Db 305 GACACCAACGAAAGTGAATAAAGTCAAAATCAAGAGCCACTCTCTAAGAAAGATTTCTCG 364
QY 121 CGGTTTGTCACTTTTCCAATCAAGTACGCTGATATTTGGAAAAATGATATAAACAGGCACAG 180
Db 365 CGGTTTGTCACTTTTCCAATCAAGTACGCTGATATTTGGAAAAATGATATAAACAGGCACAG 424
QY 181 GCTTCCCTCTGGACAGCAGAGAGGTCGACTTATCAAGAGGATCTCCCTCACTGGAAACAG 240
Db 425 GCTTCCCTCTGGACAGCAGAGAGGTCGACTTATCAAGAGGATCTCCCTCACTGGAAACAG 484
QY 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCACTCTTAGCCTTTTTCAGCCAGTGTAT 300
Db 485 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCACTCTTAGCCTTTTTCAGCCAGTGTAT 544
QY 301 GGAATTTGTAATGAAAAATTTGGTGAGCGCTTTTAGTCAGGAGGTGACGGTTCAGAGGCT 360
Db 545 GGAATTTGTAATGAAAAATTTGGTGAGCGCTTTTAGTCAGGAGGTGACGGTTCAGAGGCT 604
QY 361 CGCTGTTTCTATGCTTCAAAATTTCTCATCAGAGATGTTCACTCAGAGATGACAGTTTG 420
Db 605 CGCTGTTTCTATGCTTCAAAATTTCTCATCAGAGATGTTCACTCAGAGATGACAGTTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTTAATGCAATTTGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTTAATGCAATTTGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGGCATGGATAGCAGATAGAAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGGCATGGATAGCAGATAGAAAA 784
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCCCTTTCGCTGCTGAGAGGATTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGGAAGAGTGGTGGCCCTTTCGCTGCTGAGAGGATTTTCTTCTCAGGA 844
QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Db 845 TCTTTTGTGCTATATTTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 904
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGGCTGCTGATTTCCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGGCTGCTGATTTCCAA 964
QY 721 TACTTAGTAATAAGCCCTCAGAAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA 780
Db 965 TACTTAGTAATAAGCCCTCAGAAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA 1024
QY 781 ATTGAGCAGAGTTTTAAACAGAGCCCTTGGCCAGTTGGCCCTCATTTGGAATGAATTCAT 840
Db 1025 ATTGAGCAGAGTTTTAAACAGAGCCCTTGGCCAGTTGGCCCTCATTTGGAATGAATTCAT 1084
QY 841 TTGATGAAACAGTACATTTGAGTGTGATGTCAGCAGATTTACTTTGTGGAACCTTGGATTTCTCA 900
Db 1085 TTGATGAAACAGTACATTTGAGTGTGATGTCAGCAGATTTACTTTGTGGAACCTTGGATTTCTCA 1144
QY 901 AAGTTTTTTCAGGAGAAAAATCCCTTTTGTATTTATGGAAGAAACATTTCTTTAGAGGAAAA 960
Db 1145 AAGTTTTTTCAGGAGAAAAATCCCTTTTGTATTTATGGAAGAAACATTTCTTTAGAGGAAAA 1204
QY 961 ACAAAATTTCTTTGAGAACAGGATTTTCAGAGTATCAGCGCTTTTGGCAGTTATGGCAGAAC 1020

Db 1205 ACAAAATCTTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGAGAAACC 1264
Qy 1021 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1053
Db 1265 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1297

RESULT 3

US-10-019-733-12
; Sequence 12, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-12

Query Match 99.8%; Score 1051.4; DB 4; Length 1053;
Best Local Similarity 99.9%; Pred. No. 2.7e-310;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGGCGACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGGCGACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Qy 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTGC 120
Db 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTGC 120
Qy 121 CGGTTTGTCTATCTTCCATCCAGTACCTCATATTTGGAATAATGTATAACAGGCACAG 180
Db 121 CGGTTTGTCTATCTTCCATCCAGTACCTCATATTTGGAATAATGTATAACAGGCACAG 180
Qy 181 GCTTCCTTCTGGACAGCAGAGAGGTGCACTTATCAAGAGTCTCCCTCACTGGAACAG 240
Db 181 GCTTCCTTCTGGACAGCAGAGAGGTGCACTTATCAAGAGTCTCCCTCACTGGAACAG 240
Qy 241 CTTAAGCAGATGAGAAAGTACTTCTCTCACAATCTTAGCTTTTGGACGCGAGTAT 300
Db 241 CTTAAGCAGATGAGAAAGTACTTCTCTCACAATCTTAGCTTTTGGACGCGAGTAT 300
Qy 301 GGAATGTAAATGAAATTTGGTGGAGCGCTTTTATGTCAGGAGGTGCAGGTTCCAGAGCT 360
Db 301 GGAATGTAAATGAAATTTGGTGGAGCGCTTTTATGTCAGGAGGTGCAGGTTCCAGAGCT 360
Qy 361 CGCTGTTTCTATGCTTTCAAAATCTCATCAGAAATGTTCACTCAGAGATGTCAGTTG 420
Db 361 CGCTGTTTCTATGCTTTCAAAATCTCATCAGAAATGTTCACTCAGAGATGTCAGTTG 420
Qy 421 CTGATAGACCTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 421 CTGATAGACCTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Qy 481 ACCATGCCCTTATGTTAAGAAAAAGCAGATTTGGCGCTTTCGATGGATGACAGATAGAAA 540
Db 481 ACCATGCCCTTATGTTAAGAAAAAGCAGATTTGGCGCTTTCGATGGATGACAGATAGAAA 540

Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTGCTGTAGAAAGAGTTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTGCTGTAGAAAGAGTTTTCTTCTCAGGA 600
Qy 601 TCTTTTGGCTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGTCAGGATCTCATCTTTTCC 660
Db 601 TCTTTTGGCTCTATATTTCTGGCTAAAGAGAGAGAGGTCTTATGTCAGGATCTCATCTTTTCC 660
Qy 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCTCTGACTTTTGTGCTTGGCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCTCTGACTTTTGTGCTTGGCTGATGTTCCAA 720
Qy 721 TACTTAGTAAATAAGCCTTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGTCAA 780
Db 721 TACTTAGTAAATAAGCCTTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGTCAA 780
Qy 781 ATTGAGCAGGAGTTTTTAAACAGAAAGCTTTCAGAGTGGCTCATTTGGAATGAATTCATT 840
Db 781 ATTGAGCAGGAGTTTTTAAACAGAAAGCTTTCAGAGTGGCTCATTTGGAATGAATTCATT 840
Qy 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTTGTTGGAATTCATCTTCA 900
Db 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTTGTTGGAATTCATCTTCA 900
Qy 901 AAGGTTTTTTCAGGCGAGAAATCTTTTGAATTTTATGAAAAACATTTCTTTAGAGGAAA 960
Db 901 AAGGTTTTTTCAGGCGAGAAATCTTTTGAATTTTATGAAAAACATTTCTTTAGAGGAAA 960
Qy 961 ACAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
Db 961 ACAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
Qy 1021 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1053
Db 1021 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1053

RESULT 4

US-10-019-733-4
; Sequence 4, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-4

Query Match 99.8%; Score 1051.4; DB 4; Length 1081;
Best Local Similarity 99.9%; Pred. No. 2.7e-310;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGGCGACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 20 ATGGGCGACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 79
Qy 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGC 120
Db 80 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGC 139

Qy 121 CGGTTTGTCATCTTTCCAAATCCAGTACCCTGATATTTGGAAAAATGTATAACAGGCACAG 180
Db |||||
Qy 140 CGGTTTGTCATCTTTCCAAATCCAGTACCCTGATATTTGGAAAAATGTATAACAGGCACAG 199
Db |||||
Qy 181 GCTTCCCTTCTGGACAGCAGAAAGAGTGACTTATCAAGAGATCTCCCTCACTGAGAAACAG 240
Db |||||
Qy 200 GCTTCCCTTCTGGACAGCAGAAAGAGTGACTTATCAAGAGATCTCCCTCACTGAGAAACAG 259
Db |||||
Qy 241 CTTAAACAGATGAGAGTACTTCTCTCTCACATCTTAGCCTTTTGGCAGGCAGTGTAT 300
Db |||||
Qy 260 CTTAAACAGATGAGAGTACTTCTCTCTCACATCTTAGCCTTTTGGCAGGCAGTGTAT 319
Db |||||
Qy 301 GGAATTGTAATGAAAATTTGGTGAGCGCTTTAGTCAGAGAGTGCAGGTTCCAGAGGCT 360
Db |||||
Qy 320 GGAATTGTAATGAAAATTTGGTGAGCGCTTTAGTCAGAGAGTGCAGGTTCCAGAGGCT 379
Db |||||
Qy 361 CGCTGTTTCTATGCTTTCAAATCTCATCAGAAATGTTCTCATCAGAGATGACAGTTTG 420
Db |||||
Qy 380 CGCTGTTTCTATGCTTTCAAATCTCATCAGAAATGTTCTCATCAGAGATGACAGTTTG 439
Db |||||
Qy 421 CTGATAGACATTCATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATGAA 480
Db |||||
Qy 440 CTGATAGACATTCATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATGAA 499
Db |||||
Qy 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTGGCATGGATAGCAGATAGAAAA 540
Db |||||
Qy 500 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTGGCATGGATAGCAGATAGAAAA 559
Db |||||
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTCTGCTGTAGAGAGTTTCTTCTCAGGA 600
Db |||||
Qy 560 TCTACTTTTGGGAAAGAGTGGTGGCTTTCTGCTGTAGAGAGTTTCTTCTCAGGA 619
Db |||||
Qy 601 TCTTTTGTCTATATTTCTGCTTAAGAAAGAGGTCTTATGCCAGGACTCATTTTTCC 660
Db |||||
Qy 620 TCTTTTGTCTATATTTCTGCTTAAGAAAGAGGTCTTATGCCAGGACTCATTTTTCC 679
Db |||||
Qy 661 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTCATGTTCCAA 720
Db |||||
Qy 680 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTCATGTTCCAA 739
Db |||||
Qy 721 TACTTAGTAATAAGCCTTTCAGAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA 780
Db |||||
Qy 740 TACTTAGTAATAAGCCTTTCAGAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA 799
Db |||||
Qy 781 ATTGACGAGGATTTTAAACAGAGCCTTGCAGTGGCCTCATTTGGAATGATTCAT 840
Db |||||
Qy 800 ATTGACGAGGATTTTAAACAGAGCCTTGCAGTGGCCTCATTTGGAATGATTCAT 859
Db |||||
Qy 841 TTGATGAAAACAGTACATTTGAGTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATTCTCA 900
Db |||||
Qy 860 TTGATGAAAACAGTACATTTGAGTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATTCTCA 919
Db |||||
Qy 901 AAGGTTTTTTCAGGAGAAAATCCTTTTGATTTTATGAAAAACATTTCTTTAGAGAGAAA 960
Db |||||
Qy 920 AAGGTTTTTTCAGGAGAAAATCCTTTTGATTTTATGAAAAACATTTCTTTAGAGAGAAA 979
Db |||||
Qy 961 ACAAATTTCTTGAAGAACAGATTTTCAGAGATCAGCGTTTTCAGATTTATGGCAGAAACC 1020
Db |||||
Qy 980 ACAAATTTCTTGAAGAACAGATTTTCAGAGATCAGCGTTTTCAGATTTATGGCAGAAACC 1039
Db |||||
Qy 1021 ACAGATAACGCTTCACCTTGGATGACAGATTTT 1053
Db |||||
Qy 1040 ACAGATAACGCTTCACCTTGGATGACAGATTTT 1072
Db |||||

RESULT 5

US-09-962-665-9
; Sequence 9, Application US/09962665
; Patent No. 6537759
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jf., Vincent P.
; TITLE OF INVENTION: POLYLPOLYGLUTAMATE SYNTHETASE GENE SEQUENCE

; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING THE
; TITLE OF INVENTION: TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015004
; CURRENT APPLICATION NUMBER: US/09/962,665
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128,1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc_feature
; LOCATION: 1636, 1738, 2259
; OTHER INFORMATION: n = c or t
US-09-962-665-9

Query Match 55.0%; Score 579; DB 4; Length 2500;

Best Local Similarity 75.0%; Pred. No. 6.4e-166;

Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

Qy 89 ATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGCCGGTTTGTGTCATCTTTCCAAATCCAGTACC 148
Db |||||
Qy 397 AGGATGAGCCCTGCTGAGAGAAACCCCGCCGGTTTGTGTCATCTTTCCCATCGAGTACC 456
Db |||||
Qy 149 CTGATATTTGAAAAATGTATAAACAGGCA CAGGCTTCTCTTGGACAGCAGAGAGGTCG 208
Db |||||
Qy 457 ATGATATCTGGCAGATGTATAAGAGGCGAGAGGCTTCTTTTGGACCCCGAGGAGGTTG 516
Db |||||
Qy 209 ACTTATCAAGGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAGTACTTCACT 268
Db |||||
Qy 517 ACCTCTCAAGAGACATTCAGCACTGGGAATCCCTGAAACCCCGAGGAGAGATATTTATAT 576
Db |||||
Qy 269 CTCACTCTTAGCCCTTTTTCAGCCAGTGTGAATTTGTAATGAAATTTGGTGGAGC 328
Db |||||
Qy 577 CCCATGTTCTGGCTTTCTTTGCGAAGCGATGGATAGTAATGAAACCTTTGGTGGAGC 636
Db |||||
Qy 329 GCTTTAGTCAGGAGGTGCGGTTCCAGAGGCTCGCTGTTCTTCTATGGCTTTTCAAATTTCA 388
Db |||||
Qy 637 GATTAGCAAGAGTTCAGATTACAGAGCCCGCTGTTTCTATGCTTCCAAATTTGCCA 696
Db |||||
Qy 389 TCGAGATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACATTTACATCAGAGATCCCA 448
Db |||||
Qy 697 TGGAAAAACATACATTTCTGAAATGTATAGTCTTTCTATTGACACTTACATATAAAGATCCCA 756
Db |||||
Qy 449 AGAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCCCTATCTTAAAGAAAAGCAG 508
Db |||||
Qy 757 AAGAAAGGGAATTTCTCTTCAATGCCATTGAAACCGATGCTTGTCTCAAGAAAGAGGAGCAG 816
Db |||||
Qy 509 ATTGGGCTTCCGATGGATAGCAGATGAAATCTACTTTTGGGAAAGAGTGGTGGCCT 568
Db |||||

Db 1237 ACTTATGAGAAATATTTCACTGGAAGAAAGACTAACTTTTGGAGAGAGATAGGCG 1296
Qy 989 AGTATCAGCGTTTTCAGTTATGCGAGAAACCAAGATAGCTTCCCTTGGATGCGAG 1048
Db 1297 AGTATCAGAGATGGGAGTGTCAAGTCCAAGAGAAATTTTACCTTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 7

US-09-962-677-9

; Sequence 9, Application US/09962677

; Patent No. 6759200

; GENERAL INFORMATION:

; APPLICANT: Stanton, Jr., Vincent P.

; TITLE OF INVENTION: THYMIDINE PHOSPHORYLASE GENE SEQUENCE

; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING

; TITLE OF INVENTION: THE TREATMENT OF DISEASE

; FILE REFERENCE: 11926-015003

; CURRENT APPLICATION NUMBER: US/09/962,677

; CURRENT FILING DATE: 2001-09-24

; PRIOR APPLICATION NUMBER: 09/658,659

; PRIOR FILING DATE: 2000-09-08

; PRIOR APPLICATION NUMBER: 09/596,033

; PRIOR FILING DATE: 2000-06-15

; PRIOR APPLICATION NUMBER: 09/357,743

; PRIOR FILING DATE: 1999-07-20

; PRIOR APPLICATION NUMBER: 09/357,024

; PRIOR FILING DATE: 1999-07-19

; PRIOR APPLICATION NUMBER: 60/093,484

; PRIOR FILING DATE: 1998-07-20

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 9

; LENGTH: 2500

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: 128..1464

; OTHER INFORMATION: n = g or a

; NAME/KEY: misc_feature

; LOCATION: 189

; OTHER INFORMATION: n = t or g

; NAME/KEY: misc_feature

; LOCATION: 524

; OTHER INFORMATION: n = c or g

; NAME/KEY: misc_feature

; LOCATION: 1399

; OTHER INFORMATION: n = t or a

; NAME/KEY: misc_feature

; LOCATION: 1636..1738, 2259

; OTHER INFORMATION: n = c or t

; US-09-962-677-9

Query Match

Best Local Similarity 55.0%; Score 579; DB 4; Length 2500;

Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

Qy 89 ATGAGAGCACTCCTAAGAAAGAGTTCTGCGCGTTTGTATCTTTCCATCCAGTACC 148
Db 397 AGGATGAGCGCTGCTGAGAGAAAACCCCGCGCTTGTCTATCTTCCCATCCAGTACC 456
Qy 149 CTGATATTTGAAAATGTATAACAGGCACAGGCTTCTTCTGGACAGCAAGAGTCG 208
Db 457 ATGATATCTGGCAGATGTATAAGGAGAGGCTTCTTTTGGACCGCGAGGTTG 516
Qy 209 ACTTATCAAGGATCTCCCTCCTCCTGGAACAGCTTAAAGCAGATGAGAAGTACTTCT 268
Db 517 ACCTCTCNAAGGACATTCAGCACTGGGATCCCTGAAACCGAGAGAGATATTTTAT 576

RESULT 8

US-09-949-016-2025

; Sequence 2025, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: C1001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

Qy 269 CTCACATCTTAGCCCTTTTTCAGCCAGTGATGGAATGTAAATGAAAATTTGGTGAGC 328
Db 577 CCCATGTTCTGGCTTTCTTTGACGACGATGGCAATAGTAATGAAAATTTGGTGAGC 636
Qy 329 GCTTTAGTCAGAGAGTGCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAAATCTCA 388
Db 637 GATTTAGCCAAAGAGTTTCAGATTACAGAAAGCCCGCTGTTTCTATGGCTTTCCAAATTTGCCA 696
Qy 389 TCAGAAATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTTACATCAGAGATCCCA 448
Db 697 TGGAAAACATACATTTCTGAAATGTATAGTCTCTTTATTGACACTTTACATAAAGATCCCA 756
Qy 449 AGAAAAGGGAATTTTATTAAATGCAATTTGAAACCATGCCCTATGTTAAGAAAAGAGCAG 508
Db 757 AAGAAAGGGAATTTCTCTTCAATGCCATGAAACGATGCCCTGTGTCAAGAAAGAGGCA 816
Qy 509 ATTGGGCTTCGCGATGATAGACAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCT 568
Db 817 ACTGGGCTTCGCGCTGGATTGGGACAAAGAGGCTACCTATGTTGAACGTTGTAGCCT 876
Qy 569 TTGCTGCTGTAGAAGAGTCTTTCTTCAGGATCTTTTGTCTGCTGCTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGCAGTGGAAGGCAATTTCTTTTCCGGTCTTTTTCGGTCTGATATTTCTGGCTCAAGA 936
Qy 629 AGAGAGGTCTTATGCCAGGACTCACTTTTCCAATGAACTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCCCTGCCCTCAATTTCTAATGAACTTATTAGCAGAGATGAGGTT 996
Qy 689 TTCACGTGTACTTTTGTCTGCTGATGTTTCCAACTACTTAGTAATAAAGCCCTTCAGAGAGAA 748
Db 997 TACACTGTGATTTTGTCTGCTGATGTTTCAAACACCTGTTACACAAACCATCGGAGGAGA 1056
Qy 749 GGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTCAGCAGAGGATTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATAATTAATCAATGCTGTTCGGATAGAAACAGGAGTTCTCTCACTGAGCCT 1116
Qy 809 TGCCAGTTGGCCTCATTTGGAATGAATTCATTTTGCATGAAACAGTACATTTGAGTTTGTAG 868
Db 1117 TGCCTGTGAAGCTCATTGGGATGAATGCACTCTAATGAGCAATACATTTGAGTTTGTGG 1176
Qy 869 CTGACAGATTAATTTGTGGAATCTTGGAATCTCTCAAAGGTTTTCAGGCAGAAAATCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAATCTGGGTTTTCAGAGGTTTTCAGATGAGAAACCATTTTG 1236
Qy 929 ATTTTATGGAACAACTTTCTTTTAGAGGAAACAAATTTCTTTTGAAGAACAGATTTTCAG 988
Db 1237 ACTTTATGGAGATAATTTTCACTGGAAGGAAAGACTTAACCTTCTTTTGAAGAGAGTAGGCG 1296
Qy 989 AGTATCAGCGTTTTCAGTTATGCGAGAAACCAAGATTAACGCTTTCACCTTGGATGCG 1048
Db 1297 AGTATCAGAGATGGGAGTGTCAAGTCCAACAGAGAAATTTCTTTTACCTTTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 8

US-09-949-016-2025

; Sequence 2025, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: C1001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 2025
;; LENGTH: 2479
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-2025

Query Match 54.9%; Score 578.4; DB 4; Length 2479;
Best Local Similarity 75.0%; Pred. No. 9.6e-166;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

```
Qy 89 ATGAAGAGCCACTCTCTAAGAAAGAGTCTCGCCGGTGTGTGATCTTTCCAAATCCAGTACC 148
Db |||||
Qy 397 AGGATGAGCCGCTGCTGAGAGAAACCCCGCGCTTTGTCTATCTCCCATCGAGTACC 456
Db |||||
Qy 149 CTGATATTGGAATAATGATATAACAGGCAAGGCTTCTTCTGGACAGCAGAGAGGTGG 208
Db |||||
Qy 457 ATGATATCTGCAGATGATATAAGAGGCGAGAGGCTTCTTTGGACCGCGAGAGGTGG 516
Db |||||
Qy 209 ACTTATCAAGAGTCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAGTACTTCTATCT 268
Db |||||
Qy 517 ACCTCTCAAGGACATTCAGCACTGGGAATCCCTGAACCCGAGGAGAGATATTTTATAT 576
Db |||||
Qy 269 CTCACATCTTAGCCCTTTTTCAGCCAGTGAATGTAATCAAAATTTGGTGGAGC 328
Db |||||
Qy 577 CCCATGTTCTGGCTTTCTTTCAGCAAGCGATGGCATAGTAATGAANAATTTGGTGGAGC 636
Db |||||
Qy 329 GCTTTAGTTCAGGAGGTGCGAGTTCAGAGGCTCGCTGTTTCTATGGCTTTCAAAATTTCTCA 388
Db |||||
Qy 637 GATTTAGCAAGAGTTCAGATTACAGAGCCGCTGTTTCTATGGCTTCCAAATTTGCA 696
Db |||||
Qy 389 TCGAAGATGTTCACTCAGAGATGACAGTGTGCTGATAGACACTTACATCAGAGATCCCA 448
Db |||||
Qy 697 TGGAAAACATACATCTGAAATGTATAGTCTCTTATTGACACTTACATAAAGATCCCA 756
Db |||||
Qy 449 AGAAAGGGATTTTATTATCAATGGAATGGAACCAATGCCCTATGTTAAGAAAAGAGCAG 508
Db |||||
Qy 757 AAGAAAGGGAAATTTCTCTTCAATGCCATTTGAACCGATGCCCTGTGTCTCAAGAGAGGCGAG 816
Db |||||
Qy 509 ATTGGGCTTCGAGTGAATAGACATAGAAATCTACTTTTGGGAAAGAGTGGTGGCCT 568
Db |||||
Qy 817 ACTGGGCTTCGCTGATGAGGGAAGAGGCTACTATGTTGGAAGCTGTTGAGGCT 876
Db |||||
Qy 569 TTGCTGCTGAGAGGAGTCTTCTTCTCAGGATCTTTGCTGCTATATTTCTGGCTAAAGA 628
Db |||||
Qy 877 TTGCTGCAAGTGAAGGCAATTTCTTTCCGGTCTTTTGGCTGATATTTCTGGCTCAAGA 936
Db |||||
Qy 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCAAATGAACCTCATCAGCAGAGATGAAGGAC 688
Db |||||
Qy 937 AACGAGGACTGATGCCCTGGCTCACAATTTCTAATGAACCTTATTAGCAGAGATGAGGTT 996
Db |||||
Qy 689 TTCACTGTGCTTGTCTGCTGATGTTCCAACTACTAGTAAATAAGCCTTCAGAGAAA 748
Db |||||
Qy 997 TACACTGTGATTTTGTCTGCTGATGTTCAACACCTGGTACACAAACCATCGGAGGAGA 1056
Db |||||
Qy 749 GGGTCAGGAGATCATTTGTTGATCTCTCAAAAATGAGCAGGAGTTTTAAACAGAGCCT 808
Db |||||
Qy 1057 GAGTAAAGAAATAATTTATCAATGCTGTTCGATAGAACAGAGTTCCTCACTAGGCT 1116
Db |||||
Qy 809 TGCAGTTGGCCTCATTTGGAAATGAATTTGCAATTTTGTGAAACAGTACATTCAGTTTGPAG 868
Db |||||
Qy 1117 TGCCTGTGAAGCTCATTTGGGATGAATTCGACTCTAATGAAGCAATACATTCAGTTTGTGG 1176
Db |||||
Qy 869 CTGACATTTACTTGTGGAATCTGATCTCAAGGTTTTCAGCAGCAAAATTCCTTTTG 928
Db |||||
Qy 1177 CAGACAGACTTATGCTGGAATCGGTTTTCAGCAAGGTTTTCAGAGTAGAGAACCATTTG 1236
Db |||||
Qy 929 ATTTTATGGAACCATTTCTTTAAGAGAAAACAAATTTCTTTTGAAGAAACGAGTTTCAG 988
Db |||||
Qy 1237 ACTTTATGGAGATATTTCTCTGAGAGAAAGACTAATCTTTTGGAGAGAGATAGGCG 1296
Db |||||
```

```
Qy 989 AGTATCAGCGTTTTCAGTATTGCGAGAAACACAGATAACGTTCTTCCCTTGGATGCG 1048
Db |||||
Qy 1297 AGTATCAGAGAGTGGAGTGTCAAGTCCAACAGAGATTCTTTTACCTTGGATGCTG 1356
Db |||||
Qy 1049 ATTT 1052
Db |||||
Qy 1357 ACTT 1360
Db |||||
```

RESULT 9

US-09-023-655-1370
; Sequence 1370, Application US/09023655
; Patent No. 6607879

;; GENERAL INFORMATION:
;; APPLICANT: Cocks, Benjamin G.
;; APPLICANT: Susan G. Stuart
;; APPLICANT: Jeffrey J. Seilhamer
;; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
;; TITLE OF INVENTION: EXPRESSION
;; NUMBER OF SEQUENCES: 1508

;; CORRESPONDENCE ADDRESS:
;; ADDRESSER: INCYTE PHARMACEUTICALS, INC.
;; STREET: 3174 PORTER DRIVE
;; CITY: PALO ALTO

;; STATE: CALIFORNIA
;; COUNTRY: USA
;; ZIP: 94304

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk

;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/023,655

;; FILING DATE: HEREWITH

;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER:
;; FILING DATE:

;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:

;; NAME: Zeller, Karen J.

;; REGISTRATION NUMBER: 37,071

;; REFERENCE/DOCKET NUMBER: PA-0001 US

;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (650) 855-0555

;; TELEFAX: (650) 845-4166

;; INFORMATION FOR SEQ ID NO: 1370:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 2500 base pairs

;; TYPE: nucleic acid

;; STRANDEDNESS: single

;; TOPOLOGY: linear

;; IMMEDIATE SOURCE:

;; LIBRARY: GENBANK

;; CLONE: 936154

US-09-023-655-1370

Query Match 54.9%; Score 578.4; DB 4; Length 2500;
Best Local Similarity 75.0%; Pred. No. 9.7e-166;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

```
Qy 89 ATGAAGAGCCACTCTCTAAGAAAGAGTCTTCGCCGGTGTGTGATCTTTCCAAATCCAGTACC 148
Db |||||
Qy 397 AGGATGAGCCGCTGCTGAGAGAAACCCCGCGCTTTGTCTATCTTTCCCATCGAGTACC 456
Db |||||
Qy 149 CTGATATTGGAATAATGATATAACAGGCAAGGCTTCTTCTGGACAGCAGAGAGGTGG 208
Db |||||
Qy 457 ATGATATCTGGCAGATGATATAAGAGGCGAGAGGCTTCTTTTGGACCCGCGAGAGGTGG 516
Db |||||
Qy 209 ACTTATCAAGAGTCTCCCTCACTGGAACACAGCTTAAAGCAGATGAGAGTACTTCTATCT 268
Db |||||
```

Db 517 ACCTCTCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGATATTTTATAT 576
Qy 269 CTCACTCTTAGCCTTTTTCAGCCAGTGTGAAATTTGTAATGAAATTTGCTGGAGC 328
Db 577 CCATGTTCTGGCTTTCTTTCAGCAAGCATGATGATTAATGAAATTTGCTGGAGC 636
Qy 329 GCTTTAGTCAGGAGTTCAGGTTCCAGAGGCTCGCTGTTCTATGGCTTTCAAAATTTCA 388
Db 637 GATTTAGCCAGAGTTTCAGATTACAGAGCCGCTGTTCTATGGCTTCCAAATTTCCA 696
Qy 389 TCGAATGTTCACTCAGAGATGTACAGTTTGTGATAGACACTATCAGAGATCCCA 448
Db 697 TGGAAACATACATTCCTGAAATGTATGCTCTTCTATTTGACACTTACATTAAGATCCCA 756
Qy 449 AGAAAGGGAAATTTTATTAATCAATTTGAAACCATGCTGCTATGTTAAGAAAGGAGC 508
Db 757 AAGAAAGGAAATTTCTCTTCAATGCCATTGAAACGATGCTTGTCTAAGAAAGGAGC 816
Qy 509 ATTGGGCTTTGCGATGATAGCAGATGAAATCTACTTTTGGGAAAGAGTGTGGCCT 568
Db 817 ACTGGGCTTTGCGCTGATTTGGGACAAAGAGGCTACCTATGTTGAACGTTGTAGCCT 876
Qy 569 TTGCTGCTGATAGAGGATTTCTTCTCAGAGATCTTTGCTGCTATATTTCTGCTAAAGA 628
Db 877 TTGCTGAGTGGAAGGATTTCTTTTTCGGTTCGATATTTCTGGCTCAAGA 936
Qy 629 AGAGAGTCTTATGSCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATCAAGGAC 688
Db 937 AACGAGACTGATGCCCTGACATTTTCTAATGAACCTTATTAGCAGAGTGAAGGTT 996
Qy 689 TTCACTGTGACTTTGCTGCTGATGTTTCAATGTTTAGTAAATGAGCCTTCAGAAGAAA 748
Db 997 TACACTGTGATTTTGTCTGCTGATGTTTCAACACCTGTTACACCAACCTCGAGGAGA 1056
Qy 749 GGGTCAGGAGATCAATTTGTGATGCTGTCAAAATGAGTGTGAGTAAAGAGCCTTTCA 808
Db 1057 GAGTAAGAGAAATTAATTAATCAATGCTGTTGCGATAGAACAGGAGTTCTCACTGAGGCT 1116
Qy 809 TGGCAGTTGGCCTCATTTGGAATGAAATTTGATGAAACATACATTTGAGTTTGTAG 868
Db 1117 TGCCTGTGAAGCTCATTTGGGATGAAATTTGCACTCTAAATGAAGCAATACATTTGAGTTTGG 1176
Qy 869 CTGACAGATTAATTTGCGAATCTGAAATCTCAAAAGTGTTCAGGCAAGAAATCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAATCTGAGTTTACCAAGGTTTTCAGAGTAGAACCATTTG 1236
Qy 929 ATTTTATGGAACATTTCTTTTGAAGGAAACAAATTTCTTTTGAAGAACGAGTTTTCAG 988
Db 1237 ACTTTATGGGAATATTTTCACTGGAAGGAAAGACTAACTTCTTTGAGAAGAGTAGGCG 1296
Qy 989 AGTATCAGCGTTTTCAGTTATGCGAGAAACCAAGATACGTTTCACTTGGATGCG 1048
Db 1297 AGTATCAGAGGATGGGAGTGTGATGATCAAGTCCACAGAGAAATTTCTTTTACCTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 10
US-09-949-016-145
; Sequence 145, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 145
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-145

Query Match 54.9%; Score 578.4; DB 4; Length 2500;
Best Local Similarity 75.0%; Pred. No. 9.7e-166;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

Qy 89 ATGAGAGCCACTCCTAAGAAAGAGTTCTCCCGGTTTGTCTATCTTCCAAATCCAGTACC 148
Db 397 AGGATGAGCCCTCTGTAGAGAAACCCGCCCGCTTTGTCTATCTTCCCCTCGAGTACC 456
Qy 149 CTGATATTTGAAATGTATAACAGGCACAGGCTTCTCTTGGACAGCAGCAAGAGGTGCG 208
Db 457 ATGATATCTGCAGATGTATAAGAGGCGAGGCTTCTTTTGGACCCCGGAGGAGTTG 516
Qy 209 ACTTATCAAAGGATCTCCCTCACTGGAACAAAGCTTAAAGCAGATGAGAGTACTTCACT 268
Db 517 ACCTCTCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGATATTTTATAT 576
Qy 269 CTCACATCTTAGCCTTTTTCAGCCAGTGTGAAATTTGTAATGAAATTTGGTGGAGC 328
Db 577 CCCATGTTCTGGCTTTCTTTGCGCAAGCGATGGCATAGTAATGAAATCTTTGGTGGAGC 636
Qy 329 GCTTTAGTCAGAGGTGAGGCTTCCAGAGGCTCGCTGTTCTTATGCTTTCATAATTTCTCA 388
Db 637 GATTTAGCCAGAGATTCAGATTACAGAGCCGCTGTTTCTATGGCTTCCAAATTTGCCA 696
Qy 389 TCGAAGATGTTCACTCAGAGATGTACAGTTTGTGCTGATAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAACATACATTCGAAATGTATAGTCTTCTTATTTGACACTTACATTAAGATCCCA 756
Qy 449 AGAAAGGGAATTTTATTAATGCAATGTAACCATGCCCCCTATGTTAAGAAAGAGCAG 508
Db 757 AAGAAAGGGAATTTCTCTTCAATGCTTCAAGAACGATGCTTGTCTCAAGAAAGAGGAGC 816
Qy 509 ATTGGGCTTGGGATGATAGCAGATGAAATCTACTTTTGGGCAAGAGTGTGGCCT 568
Db 817 ACTGGGCTTGGGCTGGATTTGGGCAAGAGGCTACCTATGTTGAACGTTGTGTAGCCT 876
Qy 569 TTGCTGCTGTAGAGGAGTTTCTTCTCAGGATCTTTTGTGCTGCTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGCAAGTGAAGGCAATTTCTTTTCGGTCTTTTTCGGTGGATATTTCTGGCTCAAGA 936
Qy 629 AGAGAGTCTTATGCGCAGGACTCACTTTTTCGAATGAACTCATCAGCAGAGATGAGAGAC 688
Db 937 AACGAGACTGATGCTGGCCTCACTTTTCTAATGAACCTTATTAGCAGAGATGAGGTT 996
Qy 689 TTCACTGTGACTTTGCTGCTGATGTTTCAAAATGAGTGTGAGTAAAGAGCCTTCAGAAGAAA 748
Db 997 TACACTGTGATTTTGTGCTGCTGATGTTTCAAAACCTGTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATTAATTAATCAATGCTGTTGCGATAGAACAGGAGTTCTCACTGAGGCT 1116
Qy 809 TGGCAGTTGGCCTCATTTGGAATGAAATTTGATGAAACATGATGATTTGAGTTTGTAG 868
Db 1117 TGCCTGTGAAGCTCATTTGGGATGAAATTTGCACTCTAAATGAAGCAATACATTTGAGTTTGG 1176
Qy 869 CTGACAGATTAATTTGCGAATCTGAAATCTCAAAAGTGTTCAGGCAAGAAATCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAATCTGAGTTTACCAAGGTTTTCAGAGTAGAACCATTTG 1236
Qy 929 ATTTTATGGAACATTTCTTTTGAAGGAAACAAATTTCTTTTGAAGAACGAGTTTTCAG 988
Db 1237 ACTTTATGGGAATATTTTCACTGGAAGGAAAGACTAACTTCTTTGAGAAGAGTAGGCG 1296
Qy 989 AGTATCAGCGTTTTCAGTTATGCGAGAAACCAAGATACGTTTCACTTGGATGCG 1048
Db 1297 AGTATCAGAGGATGGGAGTGTGATGATCAAGTCCACAGAGAAATTTCTTTTACCTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

Db 1237 ACTTATGAGATATTTCTACCTGGAGGAAGACTAACTCTTTTGAGGAAGAGTAGCGG 1296
Qy 989 AGTATCAGCGTTTGCAGTTATGCGAGAAACACAGATAAGCTCTTCACTTGGATGCAG 1048
Db 1297 AGTATCAGAGTAGGAGTGTCAAGTCCAACAGAGAAATCTTTTACCTTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 11

US-08-905-223-125
; Sequence 125, Application US/08905223
; Patent No. 622029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 125:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 481 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: CDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Brain
; FEATURE:
; NAME/KEY: sig peptide
; LOCATION: 41..343
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 4.4
; OTHER INFORMATION: seq ISHILAPFAASDG/IV
US-08-905-223-125

Query Match 41.9%; Score 441; DB 3; Length 481;
Best Local Similarity 100.0%; Pred. No. 3.2e-124;
Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGGCGACCGGAAAGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
Db 41 ATGGGCGACCGGAAAGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA 100
Qy 61 GACACCAACGAAAGTGAATAAAGTCAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 101 GACACCAACGAAAGTGAATAAAGTCAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 160

Qy 121 CGTTTGTCACTTTTCCATCCAGTACCTCGATATTTGAAAATGTATAAACAGGCACAG 180
Db 161 CGTTTGTCACTTTTCCATCCAGTACCTCGATATTTGAAAATGTATAAACAGGCACAG 220
Qy 181 GCTTCTTCTGACAGCAGAGAGGTGCACTTATCAAGGATCTCCCTCACTGGAACAAG 240
Db 221 GCTTCTTCTGACAGCAGAGAGGTGCACTTATCAAGGATCTCCCTCACTGGAACAAG 280
Qy 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTTCACTTCTAGCCCTTTTTCAGCCAGTGAT 300
Db 281 CTTAAAGCAGATGAGAAGTACTTCTCATCTTCACTTCTAGCCCTTTTTCAGCCAGTGAT 340
Qy 301 GGAATTGTAATGAAAATTTTGGTGGAGCGCTTTAGTCAGAGGTGCGGTTCAGAGGCT 360
Db 341 GGAATTGTAATGAAAATTTTGGTGGAGCGCTTTAGTCAGAGGTGCGGTTCAGAGGCT 400
Qy 361 CGCTGTTCTATGCGCTTTTCAAAATTCATCGAAGATGTTCACTCAGAGATGACAGTTTG 420
Db 401 CGCTGTTCTATGCGCTTTTCAAAATTCATCGAAGATGTTCACTCAGAGATGACAGTTTG 460
Qy 421 CTGATAGACACTTACATCAGA 441
Db 461 CTGATAGACACTTACATCAGA 481

RESULT 12

US-08-307-499-1/c
; Sequence 1, Application US/08307499
; Patent No. 5651972
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.U.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,499
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF35-1.FWCCI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:


```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 138..1460
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2456..2659
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3070..3330
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3356..4180
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4392..5894
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6171..6398
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6447..6875
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6928..7431
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 7454..7858
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 7895..8155
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 8215..8682
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 8715..9539
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 9562..10272
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 10316..11908
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 11971..12780
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 12829..13107
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 13149..14171
; US-08-307-499-14

Query Match 40.8%; Score 429.8; DB 1; Length 14176;
Best Local Similarity 66.2%; Pred. No. 5.9e-120;
Matches 651; Conservative 0; Mismatches 327; Indels 5; Gaps 2;

QY 71 AAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTCTCGCGGTTTGTC 130
Db 3336 AAAATGATTTTATACAGATGAGGCTATCTCTCAAGATCTGATCTAGGTTGTTA 3395

QY 131 TCTTTCCAAATCCAGTACCTGATATTTGGAATAATGTATPAAACAGGACAGGCTTCTTCT 190
Db 3396 TTTTCCCTATTAAAGTATCATGATATCTGGAATAATGTATPAAACATCATGTCGGAAGTTT 3455

QY 191 GGACGACGAGAGGTGCGACTTATCAAGAGATCTCCCTCACTGGNACAGCTTAAAGCAG 250

```

```

RESULT 14
US-09-299-268-1/c
; Sequence 1, Application US/09299268
; Patent No. 6217882
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi ueia, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Foxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik

```

```

Db 3456 GGACCGTTGAAGAGTAGATTTATCAAAAGATTTAGATGATGGGATAAATTAACATAAG 3515
QY 251 ATGAGAAGTACTTCTCATCTCTCCTCCTAGCCCTTTTTCAGCCAGTGTGGAATTTGTA 310
Db 3516 ACGAAAAATACTTTTATAAAACATATACTAGCATTTTTCATCTAGTGTGTAATTTGTA 3575
QY 311 ATGAAAAATTTGGTGAGCGCTTTAGTTCAGGAGGTGCAGGTTCCAGAGGCTCGCTGTTCT 370
Db 3576 ATGAGAAATTTAGCGGAAAGATTTTATGTGATGACAGTGTTCAGAGGACGATGTTCT 3635
QY 371 ATGGCTTTCAAATCTCATGAGAAATGTTCACTCAGAGATGTACAGTGTTCGCTGATGACA 430
Db 3636 ATGGATTTCAAATAGCTATGGAATAATTCATTCAAGAAATGTATAGTATTTATTAATAGA 3695
QY 431 CTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAACCATGCCCT 490
Db 3696 CATATGTAGAGATAATATAGAAAAATGTCATTTATTTAACGCTATAGAAACAATGGAAT 3755
QY 491 ATGTTAAGAAAAAGCAGATTCGGGCTTCGATGATGATAGCAGATAGAAAAATCTACTTTTG 550
Db 3756 GCGTAAAAAGAAAGCTGATTCGGGCGAGAAATGGATATC--TAGCAACAGGTATATG 3812
QY 551 GGGAAAGAGTGGTGGCCTTTGCTGCTGTAGAAAGAGTGTTCCTCAGGATCTTTTGGCTG 610
Db 3813 GAGAAAGAGTAGTAGCATTTTCAGCTGTGGAGGGAATATTCCTTCTGCTTCAATTTGCTG 3872
QY 611 CTATATCTGCTAAAGAAAGAGAGTCTTATGCGAGGACTCACTTTTCCAAATGAATCA 670
Db 3873 CTATATTTTGGATAAAAAAGCGAGGATTCGATCCCGGATTAACATTTTCTAATGAACATA 3932
QY 671 TCAGCAGAGATGAAAGACTTCCTCCTGACTTTGCTTGCCTGATGTTCCAAATCTTAGTAA 730
Db 3933 TAAGTAGAGACGAAGGTTTACATTTGCTGTTTTCGCTGTAAAGTTTAAACATTTATTTAC 3992
QY 731 ATAAGCCTTCAGAAAGAGGCTCAGGAGATCAITTTGATGCTGTCAAAATTTGACAGG 790
Db 3993 ATCCACCATCTAAGGAAGTTATAAGTCGATATCATTTGATGCGGTATATAGAAAGG 4052
QY 791 AGTTTTTAAAGAGCCTTTCAGGTTGGCTCAITTTGGAATGAATTTGATGATGAAAC 850
Db 4053 AGTTTTTACAGTGTCTATTCGGTGGATCTTATAGGTATGAATTTGTTGTTAATGCTC 4112
QY 851 AGTACATTTGATTTGTAGCTGACATTTACTTGTGGAATCTGATCTCAAGAGTTTTC 910
Db 4113 AGTATATAGAAATTCGTCAGATAGATTAATTAACAGAGTTAGGTTGTGAAAG--TCTCA 4170
QY 911 AGCGAGAAATCCTTTTTCATTTTATGGAATAATTTCTTTAGAGGAAAAACAAATTTCT 970
Db 4171 ATGTATATATCTTTTAGCTTTTATGGAGTATATATCACTAGAGTATAGACTAATTTT 4230
QY 971 TTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTTATGGCAGAAACACAGATPACG 1030
Db 4231 TCGRACGAGGATTTAGTCAATATCAAAAGATCGGGGTGTTTACAAATTAAGAGAGAATA 4290
QY 1031 TCTTCACCTTGGAGCAGATTTT 1053
Db 4291 TATTTTCTCGGATATAGATTTT 4313

```

STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.

ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: /09/299, 268
APPLICATION NUMBER: US/09/299, 268
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/901,127
FILING DATE:
APPLICATION NUMBER: US 07/908, 241
FILING DATE: 1-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908, 630
FILING DATE: 29-JUN-1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/342, 212
FILING DATE: 21-APR-1992
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Salwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 14176 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 3852..4226
FEATURE:
NAME/KEY: CDS
LOCATION: 4585..4887
FEATURE:
NAME/KEY: CDS
LOCATION: 5131..5310
FEATURE:
NAME/KEY: CDS
LOCATION: 5760..5912
FEATURE:
NAME/KEY: CDS
LOCATION: 6786..7130
FEATURE:
NAME/KEY: CDS
LOCATION: 10148..10513

Query Match 40.8%; Score 429.8; DB 3; Length 14176;
Best Local Similarity 66.2%; Pred. No. 5.9e-120;
Matches 651; Conservative 0; Mismatches 327; Indels 5; Gaps 2;

Qy 71 AAATGAAATAAGTCAAATGAAGAGCCACTCTCTAAGAAAGAGTCTCGCGGGTTTGTCA 130
Db 10841 AAATGATTTTATACAGATGAGGCTATCTTCAAGAGTCTGATCTAGGTTCGTTA 10782

Qy 131 TCTTCCAAATCCAGTACCTGATATTTGGAAATGATATAACAGGCAAGGCTTCCTTCT 190
Db 10781 TTTTCCCTATTAAGTATCATGATATCTGGAATATGATATAAACAATCATGTCGAAGTTT 10722

Qy 191 GGCAGCAGAGAGAGTCTGACTTATCAAGGATCTCCCTCACTGGAAACAAGCTTAAAGCAG 250
Db 10721 GGCAGTGAAGAGTAGATTTATCAAAAGATTTAGATGATGGGATAAATTAACATAAG 10662

Qy 251 ATGAGAAAGTACTTCTCATCTCTCAATCTTAGCCTTTTGGAGCCAGTGAATGTAA 310
Db 10661 ACAGAAATATCTTTATATAAACAATATCTAGCATTTTGGCATCTAGTGTATGTAA 10602

Qy 311 ATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGTGGAGTTCCAGAGCTCGCTCTTCT 370
Db 10601 ATGAGAAATTTAGCGGAAAGATTTTATGTGGATGTACAGTGTTCAGAGGCAAGATGTTCT 10542

Qy 371 ATGCTTTTCAAAATCTCATCGAGATGTTCCTCAGAGATGTACAGTTTGTGATAGACA 430
Db 10541 ATGGAATTTCAAAATAGCTATGGAAATATTCATTCAAGAAATGATAGTTTATTAATAGATA 10482

Qy 431 CTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAAACCATGCCCT 490
Db 10481 CATATGTAAGAGATAATATAGAAAAAATGCAATTTTATTAACGCTATAGAAACAATGGAAT 10422

Qy 491 ATGTTAAGAAAAAAGCAGATTTGGGCTTGGATGTAGATAGCAGATAGAAAAATCTACTTTTG 550
Db 10421 GCGTAAAAAAGAAAGCTGATTTGGGCCAGAAATGATATC--TAGCAACAGGATATATG 10365

Qy 551 GGGAAAGAGTGGTGGCCCTTTGCTCTGTAGAGAGATTTTCTCTCAGAGATCTTTTGTCTG 610
Db 10364 GAGAAAGAGTAGTAGCATTTGCAGCTGTGGAGGAATATCTTTTCTGGTTCAITTTGCTG 10305

Qy 611 CTATATTTCTGGCTAAAGAGAGAGGTCTTATGCGAGACTCACTTTTCCAAATGAACCTCA 670
Db 10304 CTATATTTTGGATAAAAAAAGCAGGATTTGATGCGCGGATTAACATTTTCTAATGAACCTAA 10245

Qy 671 TCAGCAGAGATGAAGGACTTCACCTGTGCTTGGCTTGCCTGATGTTCCAAATCTAGTAA 730
Db 10244 TAAGTAGAGAGGAGGTTTACATTTTGGGTGTTTAATGTTTAAACATTTATTATAC 10185

Qy 731 ATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGTCAAAATTTGAGCAGG 790
Db 10184 ATCCACCATCTAGGAAGTTTATAACGTCGATATCATTTGATCGGTTTATATAGAAAGG 10125

Qy 791 AGTTTTTAACAGAGCCTTGCCAGTTGGCCCTCATTTGGAATGAATTTGATGAAAC 850
Db 10124 AGTTTTTGACAGTTGCTATTTCCGGTGGATCTTATAGGTATGAAATTTGTTTAAATGTCTC 10065

Qy 851 AGTACATTTGATTTGTAGCTGACAGATTTCTTGGAACTTGGATTTCAAGGTTTTTC 910
Db 10064 AGTATATAGAAATTCGTCGAGATAGTTTAAACAGAGTTAGGTTTGTGANAAG--TCTCA 10007

Qy 911 AGGCAGAAATCTCTTTTGTATTTTATGAAACAATTTCTTTAGAAAGAAAAACAATTTCT 970
Db 10006 ATGTATATATCTTTTGTAGCTTTTATGGATATATCATAGAGGTAAGACTAATTTT 9947

Qy 971 TTGAGAAACGAGTTTCAAGATATCAGCGTTTTCAGTTATGCGCAGAAACCAAGATAAG 1030
Db 9946 TCGAACGACGAGTTAGTGAATATCAAAAGATGGGGTGTTCACAAATAAGAGAGATA 9887

Qy 1031 TCTTCACCTTGGATGAGATTTT 1053
Db 9886 TATTTTCTACGGATATAGATTTT 9864

RESULT 15
US-09-299-268-14
; Sequence 14, Application US/09299268
; Patent No. 6217882
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:

ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/09/299,268
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/901,127
FILING DATE:
APPLICATION NUMBER: US 07/908,241
FILING DATE: 1-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,630
FILING DATE: 29-JUN-1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/342,212
FILING DATE: 21-APR-1992
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UP35.1.FWCCI
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 14176 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 138..1460
FEATURE:
NAME/KEY: CDS
LOCATION: 2456..2659
FEATURE:
NAME/KEY: CDS
LOCATION: 2809..3030
FEATURE:
NAME/KEY: CDS
LOCATION: 3070..3330
FEATURE:
NAME/KEY: CDS
LOCATION: 3356..4180
FEATURE:
NAME/KEY: CDS
LOCATION: 4392..5894
FEATURE:
NAME/KEY: CDS
LOCATION: 6171..6398
FEATURE:
NAME/KEY: CDS
LOCATION: 6447..6875
FEATURE:
NAME/KEY: CDS
LOCATION: 6928..7431
FEATURE:
NAME/KEY: CDS
LOCATION: 7454..7858
FEATURE:

NAME/KEY: CDS
FEATURE: 7895..8155
NAME/KEY: CDS
LOCATION: 8215..8682
NAME/KEY: CDS
LOCATION: 8715..9539
NAME/KEY: CDS
LOCATION: 9562..10272
NAME/KEY: CDS
LOCATION: 10316..11908
NAME/KEY: CDS
LOCATION: 11971..12780
NAME/KEY: CDS
LOCATION: 12829..13107
NAME/KEY: CDS
LOCATION: 13149..14171
US-09-299-268-14
Query Match 40.8%; Score 429.8; DB 3; Length 14176;
Best Local Similarity 66.2%; Pred. No. 5.9e-120;
Matches 651; Conservative 0; Mismatches 327; Indels 5; Gaps 2;
QY 71 AAAGTGAATAAAGTCAATGAAGAGCACCTCCTAAGAAAGAGTTCTCGCCGGTTTGTCA 130
Db 3336 AAAATGATTTTATACACGAATGGAGCTATTTCTCAAGAGTCTGATTTCTAGGTTGTTA 3395
QY 131 TCTTTCCAATCCAGTACCTCGATATTTGGAAAAATGTATAAAGGACAGGCTTCCTTCT 190
Db 3396 TTTTCCCTATTAAAGTATCATGATATCTGAAAAATGTATAAACAATCAGTGCGCAAGTTT 3455
QY 191 GGACAGCAGAGAGGTGCACTTATCAAGAGGATCTCCCTCACTGGAAACAGCTTAAGCAG 250
Db 3456 GGACCGTTGAAGAGAGTAGATTTATCAAAAGATTTAGATGATTTGGGATAAATTAACATAAG 3515
QY 251 ATGAGAAGTACTTCACTCTCACATCTTAGCCTTTTTCAGCCAGTGTATGAAATTTGTA 310
Db 3516 ACGAAATATCTTTATAAACAATATCTAGCAATTTTTCATCTAGTGTATTTGTA 3575
QY 311 ATGAAAAATTTGGTGAGCGCTTTTAGTCAGAGGTGCAGGTTCCAGAGGCTCGCTGTTCT 370
Db 3576 ATGAGAAATTTAGCGGAAAGATTTTATGTGGATGTACAGTGTTCAGAGGCAAGATGTTCT 3635
QY 371 ATGGCTTTCAAATTTCTATCGAGATGTTCTACTCAGAGATGTACAGTTTGTGATAGACA 430
Db 3636 ATGGATTTCAAATAGCTATGGAATAATTCATTCAGAAATGTATAGTTTATTAATAGATA 3695
QY 431 CTATACATCAGAGATCCCAAGAAAGGGAAATTTTATTTAAATGCAATTTGAAACCATGCCCT 490
Db 3696 CATATGTAAGAGATAATAGAAAAAATGCAATTTTATTAACGCTATAGAAACAATGGAAT 3755
QY 491 ATGTTAAGAAAAAGCAGATTTGGGCTTTGCGATGTAGTAGCAGATAGAAAAATCTACTTTTG 550
Db 3756 GCGTAAAAAAGAAAGCTGATTTGGGCCAGAAATGGATATC---TAGCAACAGGTATATG 3812
QY 551 GGGAAAGAGTGTGGCTTTTCCTGCTGTAGAGAGGATTTTCTCTCAGAGATCTTTTGTCTG 610
Db 3813 GAGAAAGAGTAGTAGCATTTTCAGCTGTGGAGGGAATATTTCTTTCTGCTTCAATTTGCTG 3872
QY 611 CTATATTTCTGCTAAAGAGAGAGTCTTATGCCAGAGTCTCACTTTTCCCAATGAACTCA 670
Db 3873 CTATATTTTGGATAAAAAAAGCAGGATTTGATGCCCGGATTAACATTTTCTAATGAACTAA 3932
QY 671 TCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGTGCTGTATGTTTCAATCTAGTAA 730
Db 3933 TAAGTAGAGCAGAGGTTTACATTTGATTTTGGGTTTGGGTTTAAATGTTTAAACATTTATAC 3992

Qy	731	ATAAGCCTTCAGAAAGAGGTCAAGGAGATCATTTGTTGATGCTGTCAAAATTGAGCAGG	790
Db	3993		4052
Qy	791	AGTTTTTAACAGAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTCATTTTGATGAAC	850
Db	4053		4112
Qy	851	AGTATATTGACAGTTGCTTATTCGGTGGATCTTATAGGTATGAATTTGTTTAAATGTCTC	910
Db	4113		4170
Qy	911	AGGAGAAATCCCTTTTGAATTTATGGAACATTTCTTTAGAAAGGAAACAAATTTCT	970
Db	4171		4230
Qy	971	TTGGAACCGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACCCAGATAACG	1030
Db	4231		4290
Qy	1031	TCTTCACCTTGGATGCAGATTTT	1053
Db	4291		4313

Search completed: September 26, 2005, 02:25:52
Job time : 233 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 26, 2005, 01:35:20 ; Search time 577 Seconds
(without alignments)
12201.951 Million cell updates/sec

Title: US-10-698-228-2
Perfect score: 1053
Sequence: 1 atgggcaccggaagcc.....tcacctggatgcagatttt 1053

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 7400704 seqs, 3343079526 residues
Total number of hits satisfying chosen parameters: 14801408

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1053	100.0	1053	17	US-10-698-228-2
2	1053	100.0	4955	18	US-10-172-118-71
3	1053	100.0	4955	18	US-10-342-887-71
4	1053	100.0	4955	18	US-10-698-228-3
5	1051.4	99.8	1053	18	US-10-698-228-12
6	1051.4	99.8	1081	18	US-10-698-228-4
7	578.4	54.9	1989	9	US-09-925-301-505

8	578.4	54.9	2216	15	US-10-084-817-342	Sequence 342, App
9	578.4	54.9	2482	22	US-10-220-335-514	Sequence 514, App
10	578.4	54.9	2500	9	US-09-954-456-724	Sequence 724, App
11	578.4	54.9	2500	9	US-09-954-456-1169	Sequence 1169, App
12	578.4	54.9	2500	9	US-09-954-456-1827	Sequence 1827, App
13	578.4	54.9	2500	18	US-10-641-643-1370	Sequence 1370, App
14	578.4	54.9	2500	20	US-10-733-878-458	Sequence 458, App
15	578.4	54.9	2500	21	US-10-843-641A-3751	Sequence 3751, App
16	578.4	54.9	2500	21	US-10-843-641A-4196	Sequence 4196, App
17	578.4	54.9	2500	21	US-10-843-641A-4854	Sequence 4854, App
18	578.4	54.9	2500	22	US-10-756-149-713	Sequence 713, App
19	570.4	54.2	2113	21	US-10-764-420-1636	Sequence 1636, App
20	569	54.0	2641	22	US-10-220-335-170	Sequence 170, App
21	561	53.3	1328	18	US-10-403-571-75	Sequence 75, App1
22	545.4	51.8	977	17	US-10-264-237-790	Sequence 790, App
23	514	48.8	1371	22	US-10-450-763-15278	Sequence 15278, App
24	447.8	42.5	186854	21	US-10-872-156-34	Sequence 34, App1
25	442	42.0	1146	9	US-09-822-830A-174	Sequence 174, App
26	432.2	41.0	1218	15	US-10-128-714-7245	Sequence 7245, App
27	403.2	38.3	963	17	US-10-369-493-34052	Sequence 34052, App
28	397	37.7	1146	15	US-10-128-714-2245	Sequence 2245, App
29	392.6	37.3	1173	17	US-10-369-493-46252	Sequence 46252, App
30	392	37.2	1242	16	US-10-032-585-6537	Sequence 6537, App
31	386.6	36.7	1292	17	US-10-369-493-27809	Sequence 27809, App
32	380.6	36.1	1248	18	US-10-424-599-126843	Sequence 126843, App
33	380.4	36.1	1200	17	US-10-369-493-45857	Sequence 45857, App
34	371.2	35.3	1361	18	US-10-424-599-40471	Sequence 40471, App
35	365.8	34.7	1586	18	US-10-424-599-73024	Sequence 73024, App
36	364.8	34.6	1206	17	US-10-320-797-2145	Sequence 2145, App
37	361.6	34.3	957	20	US-10-653-047-43	Sequence 43, App1
38	351.4	33.4	629	22	US-10-450-763-6851	Sequence 6851, App
39	349	33.1	1314	15	US-10-128-714-1245	Sequence 1245, App
40	349	33.1	1450	15	US-10-128-714-6245	Sequence 6245, App
41	349	33.1	3314	15	US-10-128-714-245	Sequence 245, App
42	349	33.1	3450	15	US-10-128-714-5245	Sequence 5245, App
43	328.2	31.2	560	9	US-09-864-761-12381	Sequence 12381, App
44	324	30.8	1169	18	US-10-425-114-31361	Sequence 31361, App
45	324	30.8	2992	20	US-10-425-115-31182	Sequence 31182, App

ALIGNMENTS

RESULT 1
US-10-698-228-2
; Sequence 2, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-2

Query Match 100.0%; Score 1053; DB 18; Length 1053;
Best Local Similarity 100.0%; Pred. No. 8.5e-295;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 ATGGCGACCGGAAAGCCGAGCGCCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db |||
Qy 1 ATGGCGACCGGAAAGCCGAGCGCCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db |||
Qy 61 GACACCAACGAAGTGAATAAGTCAAAATGAAGGCCATCTCCCTAAGAGAGTTCTCGC 120
Db |||
Qy 61 GACACCAACGAAGTGAATAAGTCAAAATGAAGGCCATCTCCCTAAGAGAGTTCTCGC 120
Db |||
Qy 121 CGGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATAACAGGCACAG 180
Db |||
Qy 121 CGGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATAACAGGCACAG 180
Db |||
Qy 181 GCTTCTCTCTGACAGCAGAGAGGTGCACTTATCAAAAGGATCTCCCTCACTGGAACAAG 240
Db |||
Qy 181 GCTTCTCTCTGACAGCAGAGAGGTGCACTTATCAAAAGGATCTCCCTCACTGGAACAAG 240
Db |||
Qy 241 CTTAAAGCAGATGAAGATCTTCATCTCTCACATCTTAGCTTTTGGAGCCAGTGAT 300
Db |||
Qy 241 CTTAAAGCAGATGAAGATCTTCATCTCTCACATCTTAGCTTTTGGAGCCAGTGAT 300
Db |||
Qy 301 GGAATGTGAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGCT 360
Db |||
Qy 301 GGAATGTGAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGCT 360
Db |||
Qy 361 CGCTGTTTCTATGCTTTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
Db |||
Qy 361 CGCTGTTTCTATGCTTTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
Db |||
Qy 421 CTGATAGACATTAATCAGAGATGCCAAGAAAAGGGAATTTTATTAAATGCAATGAA 480
Db |||
Qy 421 CTGATAGACATTAATCAGAGATGCCAAGAAAAGGGAATTTTATTAAATGCAATGAA 480
Db |||
Qy 481 ACCATGCCCTATGTTAGAAAAGAGCAGATGGGCCCTTGCATGCATGATGACGATAGAAA 540
Db |||
Qy 481 ACCATGCCCTATGTTAGAAAAGAGCAGATGGGCCCTTGCATGCATGATGACGATAGAAA 540
Db |||
Qy 541 TCTACTTTTGGGAAAGAGTGGGCTTTGCTCTGTAGAGGAGTTTCTTCTCAGGA 600
Db |||
Qy 541 TCTACTTTTGGGAAAGAGTGGGCTTTGCTCTGTAGAGGAGTTTCTTCTCAGGA 600
Db |||
Qy 601 TCTTTTGTCTGCTATTTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Db |||
Qy 601 TCTTTTGTCTGCTATTTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Db |||
Qy 661 AATGAACTCATCAGCAGATGAAGGACTTCACGTGACTTTGCTTGCCTGAGTTCCAA 720
Db |||
Qy 661 AATGAACTCATCAGCAGATGAAGGACTTCACGTGACTTTGCTTGCCTGAGTTCCAA 720
Db |||
Qy 721 TACTTAGTAAATAGCCCTTCAGAAAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAA 780
Db |||
Qy 721 TACTTAGTAAATAGCCCTTCAGAAAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAA 780
Db |||
Qy 781 ATTGAGCAGGAGTTTTTAAACAGAGCCCTTGCCAGTTGGCCCTCATTTGGAATGCAATT 840
Db |||
Qy 781 ATTGAGCAGGAGTTTTTAAACAGAGCCCTTGCCAGTTGGCCCTCATTTGGAATGCAATT 840
Db |||
Qy 841 TTGATGAAAACGATACATGATGTTGTAGCTGACAGATTAATTGTGGAATTCTTCA 900
Db |||
Qy 841 TTGATGAAAACGATACATGATGTTGTAGCTGACAGATTAATTGTGGAATTCTTCA 900
Db |||
Qy 901 AAGGTTTTTCAGGCAGAAAATCCCTTTGATTTTATGGAACAATTTCTTTAGAGGAAA 960
Db |||
Qy 901 AAGGTTTTTCAGGCAGAAAATCCCTTTGATTTTATGGAACAATTTCTTTAGAGGAAA 960
Db |||
Qy 961 ACAAAATTTCTTGAGAAAACGAGTTTCAGAGTATCAGGCTTTTGCAGTTATGGCAGAAACC 1020
Db |||
Qy 961 ACAAAATTTCTTGAGAAAACGAGTTTCAGAGTATCAGGCTTTTGCAGTTATGGCAGAAACC 1020
Db |||
Qy 1021 ACAGATAACGTTCTTCACTTTGGATGCGAGATTTT 1053
Db |||
Qy 1021 ACAGATAACGTTCTTCACTTTGGATGCGAGATTTT 1053
Db |||
```

```
RESULT 2
US-10-172-118-71
; Sequence 71, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AB036063
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-71

Query Match 100.0%; Score 1053; DB 17; Length 4955;
Best Local Similarity 100.0%; Pred. No. 2.2e-294;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGACCGGAAAGCCGAGCGCCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 245 ATGGCGACCGGAAAGCCGAGCGCGCGCTGGATCAGGATGAGAGATCATCTTCA 304

Qy 61 GACACCAACGAAGTGAATAAGTCAAAATGAAGGCCATCTCCTAAGAGAGTTCTCGC 120
Db 305 GACACCAACGAAGTGAATAAGTCAAAATGAAGGCCATCTCCTAAGAGAGTTCTCGC 364

Qy 121 CGGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATAACAGGCACAG 180
Db 365 CGGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATAACAGGCACAG 424

Qy 181 GCTTCTCTCTGACAGCAGAGAGGTGCACTTATCAAAAGGATCTCCCTCACTGGAACAAG 240
Db 425 GCTTCTCTCTGACAGCAGAGAGGTGCACTTATCAAAAGGATCTCCCTCACTGGAACAAG 484

Qy 241 CTTAAAGCAGATGAGAGTACTTTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT 300
Db 485 CTTAAAGCAGATGAGAGTACTTTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT 544

Qy 301 GGAATGTGAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGCT 360
Db 545 GGAATGTGAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGCT 604

Qy 361 CGCTGTTTCTATGCTTTTCAAAATCTCATCGAGATGTTTCACTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTTCTATGCTTTTCAAAATCTCATCGAGATGTTTCACTCAGAGATGTACAGTTTG 664

Qy 421 CTGATAGACATTAATCAGAGATGCCAAGAAAAGGGAATTTTATTAAATGCAATGAA 480
Db 665 CTGATAGACATTAATCAGAGATGCCAAGAAAAGGGAATTTTATTAAATGCAATGAA 724

Qy 481 ACCATGCCCTATGTTAGAAAACGAGTTTCAGAGTATGCGGCTTTGCGATGATGACAGATAAAA 540
Db 725 ACCATGCCCTATGTTAGAAAACGAGTTTCAGAGTATGCGGCTTTGCGATGATGACAGATAAAA 784

Qy 541 TCTACTTTTGGGAAAGAGTGGGCTTTGCTCTGTAGAGGAGTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGGGCTTTGCTCTGTAGAGGAGTTTCTTCTCAGGA 844
```

Qy 601 TCTTTTGTCTATATCTTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCATTCTTCC 660
Db 845 TCTTTTGTCTATATCTTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCATTCTTCC 904
Qy 661 AATGAACATCATCAGCAGAGATGAGGACTTCACTGTGACTTTTGGCTTGTGCTGATCTTCAA 720
Db 905 AATGAACATCATCAGCAGAGATGAGGACTTCACTGTGACTTTTGGCTTGTGCTGATCTTCAA 964
Qy 721 TACTTAGTAATAAGCCCTTCAAGAGAAAGGCTCAGGAGAGATCATTTGTGCTGTCAA 780
Db 965 TACTTAGTAATAAGCCCTTCAAGAGAAAGGCTCAGGAGAGATCATTTGTGCTGTCAA 1024
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCCTTGGCCCTCACTTGAATGAATGCAAT 840
Db 1025 ATTGAGCAGGAGTTTAAACAGAGCCCTTGGCCCTCACTTGAATGAATGCAAT 1084
Qy 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATTTGGAACTTTGGATCTTCA 900
Db 1085 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATTTGGAACTTTGGATCTTCA 1144
Qy 901 AAGTTTTTTCAGGCGAGAAATCCCTTTTGAATTTTATGAAACATTTCTTTAGAGGAAA 960
Db 1145 AAGTTTTTTCAGGCGAGAAATCCCTTTTGAATTTTATGAAACATTTCTTTAGAGGAAA 1204
Qy 961 ACAATTTTCTTTGAGAAACAGATTTTCAAGATATCAGCGTTTGTGAGTTATGGCAGAAACC 1020
Db 1205 ACAATTTTCTTTGAGAAACAGATTTTCAAGATATCAGCGTTTGTGAGTTATGGCAGAAACC 1264
Qy 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053
Db 1265 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1297

RESULT 3

US-10-342-887-71
; Sequence 71, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-71

Query Match 100.0%; Score 1053; DB 18; Length 4955;
Best Local Similarity 100.0%; Pred. No. 2.2e-294;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGGCGACCGGAAAGCGCGAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
Db 245 ATGGGCGACCGGAAAGCGCGAGCGCGGCTGGATCAGATGAGATCATCTTCA 304
Qy 61 GACACCAACCAAGTGAATAAGTCAATGAAGAGCCACTCTTCAAGAAAGATTTCTCGC 120

Db 305 GACACCAACCAAGTGAATAAGTCAATGAAGAGCCACTCTTCAAGAAAGATTTCTCGC 364
Qy 121 CGTTTGTCTATCTTCCCAATCCAGTACCTGTATTTTGGAAATGTATATAACAGGCACAG 180
Db 365 CGTTTGTCTATCTTCCCAATCCAGTACCTGTATTTTGGAAATGTATATAACAGGCACAG 424
Qy 181 GCTTCTCTTGGACAGCAGAGAGGTGACATTTATCAAGAGATCTCCCTCACTGGAAACAG 240
Db 425 GCTTCTCTTGGACAGCAGAGAGGTGACATTTATCAAGAGATCTCCCTCACTGGAAACAG 484
Qy 241 CTTAAAGCAGATGAGAGATCTTCAATCTCTCACTCTTAGCCTTTTTCGACGCAAGTAT 300
Db 485 CTTAAAGCAGATGAGAGATCTTCAATCTCTCACTCTTAGCCTTTTTCGACGCAAGTAT 544
Qy 301 GGAATGTGAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGTTCCAGAGGCT 360
Db 545 GGAATGTGAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGTTCCAGAGGCT 604
Qy 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAATGTTCACTCAGAGATGTACAGTTTG 664
Qy 421 CTGATAGACACTTATCATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 480
Db 665 CTGATAGACACTTATCATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 724
Qy 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTTCGATGATGATGATGATGATGAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTTCGATGATGATGATGATGAAA 784
Qy 541 TCTACTTTTGGGAAAGAGTGTGGCTTTTCTGCTGTAGAGGAGTTTCTTCTCAGCA 600
Db 785 TCTACTTTTGGGAAAGAGTGTGGCTTTTCTGCTGTAGAGGAGTTTCTTCTCAGCA 844
Qy 601 TCTTTTGTCTATATCTTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCATTCTTCTCC 860
Db 845 TCTTTTGTCTATATCTTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCATTCTTCTCC 904
Qy 661 AATGAACATCATCAGCAGAGATGAGGACTTCACTGTGACTTTTGGCTTGTGCTGATTTCAA 720
Db 905 AATGAACATCATCAGCAGAGATGAGGACTTCACTGTGACTTTTGGCTTGTGCTGATTTCAA 964
Qy 721 TACTTAGTAATAAGCCCTTCAAGAGAAAGGCTCAGGAGAGATCATTTGTGCTGTCAA 780
Db 965 TACTTAGTAATAAGCCCTTCAAGAGAAAGGCTCAGGAGAGATCATTTGTGCTGTCAA 1024
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCCTTGGCCCTCACTTGAATGAATGCAAT 840
Db 1025 ATTGAGCAGGAGTTTAAACAGAGCCCTTGGCCCTCACTTGAATGAATGCAAT 1084
Qy 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATTTGGAACTTTGGATCTTCA 900
Db 1085 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATTTGGAACTTTGGATCTTCA 1144
Qy 901 AAGTTTTTTCAGGCGAGAAATCCCTTTTGAATTTTATGAAACATTTCTTTAGAGGAAA 960
Db 1145 AAGTTTTTTCAGGCGAGAAATCCCTTTTGAATTTTATGAAACATTTCTTTAGAGGAAA 1204
Qy 961 ACAATTTTCTTTGAGAAACAGATTTTCAAGATATCAGCGTTTGTGAGTTATGGCAGAAACC 1020
Db 1205 ACAATTTTCTTTGAGAAACAGATTTTCAAGATATCAGCGTTTGTGAGTTATGGCAGAAACC 1264
Qy 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053
Db 1265 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1297

RESULT 4

US-10-698-228-3
; Sequence 3, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:

; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 3
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-3

Query Match 100.0%; Score 1053; DB 18; Length 4955;
Best Local Similarity 100.0%; Pred. No. 2.2e-294;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGCACCGGAAAGCGCGGCGCTGGATCAGGATGAGATCATCTTCA 60
Db |
Qy 245 ATGGCGCACCGGAAAGCGCGGCGCTGGATCAGGATGAGATCATCTTCA 304
Db |
Qy 61 GACACCAACGAAAGTAAATGAATCAAAATGAAGAGCCACTCTTAAGAAAGATTCTGC 120
Db |
Qy 305 GACACCAACGAAAGTAAATGAATCAAAATGAAGAGCCACTCTTAAGAAAGATTCTGC 364
Db |
Qy 121 CGTTTGTGATCTTCCATCCAGTACCTGATATTTGGAAATGTATAACAGGCACAG 180
Db |
Qy 365 CGTTTGTGATCTTCCATCCAGTACCTGATATTTGGAAATGTATAACAGGCACAG 424
Db |
Qy 181 GTTCTCTTCTGGACAGAGAGAGTGCATTAACAAGAGTCCCTCACTGGAACAAG 240
Db |
Qy 425 GTTCTCTTCTGGACAGAGAGAGTGCATTAACAAGAGTCCCTCACTGGAACAAG 484
Db |
Qy 241 CTTAAGCAGATGAGAAGTACTTCACTCTCACTCTTAGCTTTTGGACGCAAGTAT 300
Db |
Qy 485 CTTAAGCAGATGAGAAGTACTTCACTCTCACTCTTAGCTTTTGGACGCAAGTAT 544
Db |
Qy 301 GGAATTTGAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGCAGAGCT 360
Db |
Qy 545 GGAATTTGAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGCAGAGCT 604
Db |
Qy 361 CGCTGTTTCTATGCTTTCAAAATTTCTCATCGAGATGTTCACTCAGAGATGACAGTTG 420
Db |
Qy 605 CGCTGTTTCTATGCTTTCAAAATTTCTCATCGAGATGTTCACTCAGAGATGACAGTTG 664
Db |
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATGAA 480
Db |
Qy 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATGAA 724
Db |
Qy 481 ACCATGCCCTTANGTTAAGAAAAAGCAGATGAGGCTTGGATGATAGCAGATAGAAAA 540
Db |
Qy 725 ACCATGCCCTTANGTTAAGAAAAAGCAGATGAGGCTTGGATGATAGCAGATAGAAAA 784
Db |
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTTTCTTCTCAGGA 600
Db |
Qy 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTTTCTTCTCAGGA 844
Db |
Qy 601 TCTTTTGTGCTATATTCTGGCTAAAGAAAGAGGCTTTATGCCAGGACTCACTTTTTC 660
Db |
Qy 845 TCTTTTGTGCTATATTCTGGCTAAAGAAAGAGGCTTTATGCCAGGACTCACTTTTTC 904
Db |
Qy 661 AATGACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 720
Db |
Qy 905 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
Db |

Qy 721 TACTTAGTAAATAAGCCCTTCAAGAAAGGCGTCAGGAGATCATTTGTGATGCTGTCAA 780
Db |
Qy 965 TACTTAGTAAATAAGCCCTTCAAGAAAGGCGTCAGGAGATCATTTGTGATGCTGTCAA 1024
Db |
Qy 781 ATTGACGAGGAGTTTAAACAGAACCTTGCAGTTGCCCTCATTTGGAATGAATGCAAT 840
Db |
Qy 1025 ATTGACGAGGAGTTTAAACAGAACCTTGCAGTTGCCCTCATTTGGAATGAATGCAAT 1084
Db |
Qy 841 TTGATGAACAGTACATTTGATGTTGTAGCTGACAGATTACTTTGTGGAATGCTTCA 900
Db |
Qy 1085 TTGATGAACAGTACATTTGATGTTGTAGCTGACAGATTACTTTGTGGAATGCTTCA 1144
Db |
Qy 901 AAGGTTTTCAGGACAGAAATCCTTTGATTTTATGGAACAAATTTCTTTAGAAGGAAA 960
Db |
Qy 1145 AAGGTTTTCAGGACAGAAATCCTTTGATTTTATGGAACAAATTTCTTTAGAAGGAAA 1204
Db |
Qy 961 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGCAGAAAC 1020
Db |
Qy 1205 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGCAGAAAC 1264
Db |
Qy 1021 ACAGATACGCTTTCACCTTCGATGCGAGTTT 1053
Db |
Qy 1265 ACAGATACGCTTTCACCTTCGATGCGAGTTT 1297
Db |

RESULT 5
US-10-698-228-12
; Sequence 12, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-12

Query Match 99.8%; Score 1051.4; DB 18; Length 1053;
Best Local Similarity 99.9%; Pred. No. 2.5e-294;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGCGCACCGGAAAGCGCGGCGCTGGATCAGGATGAGATCATCTTCA 60
Db |
Qy 1 ATGGCGCACCGGAAAGCGCGGCGCTGGATCAGGATGAGATCATCTTCA 60
Db |
Qy 61 GACACCAACGAAAGTAAATGAATCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCG 120
Db |
Qy 61 GACACCAACGAAAGTAAATGAATCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCG 120
Db |
Qy 121 CGTTTGTGATCTTCCATCCAGTACCTGATATTTGGAAATGTATAAAGCAGCAG 180
Db |
Qy 121 CGTTTGTGATCTTCCATCCAGTACCTGATATTTGGAAATGTATAAAGCAGCAG 180
Db |
Qy 181 GTTCTCTTCTGGACAGAGAGAGTGCATTAACAAGAGTCCCTCACTGGAACAAG 240
Db |
Qy 181 GTTCTCTTCTGGACAGAGAGAGTGCATTAACAAGAGTCCCTCACTGGAACAAG 240
Db |

QY 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCATCTTACGCTTTTTCAGCCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCATCTTACGCTTTTTCAGCCAGTGAT 300
QY 301 GGAAATTTGAAATGAAATTTGGTGGAGGCTTTAGTTCAGGAGTGCAGGCTCAGAGGCT 360
Db 301 GGAAATTTGAAATGAAATTTGGTGGAGGCTTTAGTTCAGGAGTGCAGGCTCAGAGGCT 360
QY 361 CGCTGTTTCTAGGCTTTCAAAATTTCTCATCGAAGATGTTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTAGGCTTTCAAAATTTCTCATCGAAGATGTTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAGGAGATGAGGCTTGGCTTGGATGATAGCAATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAGGAGATGAGGCTTGGCTTGGATGATAGCAATAGAAA 540
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTAGAAGGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTAGAAGGAGTTTCTTCTCAGGA 600
QY 601 TCTTTGCTGCTATATCTGGCTAAAGAGAGAGGCTTCTATGCCAGGACTCACTTTTCC 660
Db 601 TCTTTGCTGCTATATCTGGCTAAAGAGAGAGGCTTCTATGCCAGGACTCACTTTTCC 660
QY 661 AATGAACTCATCAGCAGATGAGGACTTCACTGTGACTTTGCTTGGCTGTAGTTCCAA 720
Db 661 AATGAACTCATCAGCAGATGAGGACTTCACTGTGACTTTGCTTGGCTGTAGTTCCAA 720
QY 721 TACTTAGTAAATAGCCCTTCAAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db 721 TACTTAGTAAATAGCCCTTCAAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
QY 781 ATTGAGCAGGAGTTTAAACAGAGCTTGCAGTTGGCTCATTTGGAATGATTTGCATT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGCTTGCAGTTGGCTCATTTGGAATGATTTGCATT 840
QY 841 TTGATGAAACAGTACATTTAGTTTGTAGTACAGATTAATTTGGAATTTGGAATTTCA 900
Db 841 TTGATGAAACAGTACATTTAGTTTGTAGTACAGATTAATTTGGAATTTGGAATTTCA 900
QY 901 AAGTTTTTTCAGGAGAAATCCTTTTCAATTTTATGGAACATTTCTTTAGAGGAAA 960
Db 901 AAGTTTTTTCAGGAGAAATCCTTTTCAATTTTATGGAACATTTCTTTAGAGGAAA 960
QY 961 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTTATGCGAGAAC 1020
Db 961 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTTATGCGAGAAC 1020
QY 1021 ACAGATAAGCTTTCACCTTGGATGAGATTTT 1053
Db 1021 ACAGATAAGCTTTCACCTTGGATGAGATTTT 1053

RESULT 6

US-10-698-228-4
; Sequence 4, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06

; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-4

Query Match 99.8%; Score 1051.4; DB 18; Length 1081;
Best Local Similarity 99.9%; Pred. No. 2.5e-294;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGGCGACCCGAAAGGCCGAGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 20 ATGGGGCGACCCGAAAGGCCGAGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 79
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 80 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 139
QY 121 CGGTTTGTCACTTTTCCAATCCAGTACCCTGATATTTGGAAATGTATAAACAGCACAG 180
Db 140 CGGTTTGTCACTTTTCCAATCCAGTACCCTGATATTTGGAAATGTATAAACAGCACAG 199
QY 181 GCTTCCCTTCGGAAGAGAGGTCGACTTATCAAAAGGATCTCCCTCACTGGAACAG 240
Db 200 GCTTCCCTTCGGAAGAGAGGTTGACTTATCAAAAGGATCTCCCTCACTGGAACAG 259
QY 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCATCTTTAGCCTTTTTCAGCCAGTGAT 300
Db 260 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCATCTTTAGCCTTTTTCAGCCAGTGAT 319
QY 301 GGAAATTTGAAATGAAATTTGGTGGAGGCTTTTAGTTCAGGAGTGCAGGTTCCAGAGCT 360
Db 320 GGAAATTTGAAATGAAATTTGGTGGAGGCTTTTAGTTCAGGAGTGCAGGTTCCAGAGCT 379
QY 361 CGCTGTTTCTAGGCTTTCAAAATTTCTCATCGAAGATGTTTCACTCAGAGATGTACAGTTTG 420
Db 380 CGCTGTTTCTAGGCTTTCAAAATTTCTCATCGAAGATGTTTCACTCAGAGATGTACAGTTTG 439
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 480
Db 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 499
QY 481 ACCATGCCCTATGTTAAGAAAGGAGATGAGGCTTGGCTTGGATGATAGCAATAGAAA 540
Db 500 ACCATGCCCTATGTTAAGAAAGGAGATGAGGCTTGGCTTGGATGATAGCAATAGAAA 559
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTAGAAGGAGTTTCTTCTCAGGA 600
Db 560 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTAGAAGGAGTTTCTTCTCAGGA 619
QY 601 TCTTTTGTCTGCTATATTTCTGGCTAAAGAGAGGCTTCTATGCGAGGACTCACTTTTCC 660
Db 620 TCTTTTGTCTGCTATATTTCTGGCTAAAGAGAGGCTTCTATGCGAGGACTCACTTTTCC 679
QY 661 AATGAACTCATCAGCAGAGATGAGGACTTCACTGTGACTTTGCTTCCCTGATGTTCCAA 720
Db 680 AATGAACTCATCAGCAGAGATGAGGACTTCACTGTGACTTTGCTTCCCTGATGTTCCAA 739
QY 721 TACTTAGTAAATAGCCCTTCAAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db 740 TACTTAGTAAATAGCCCTTCAAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 799
QY 781 ATTGAGCAGGAGTTTAAACAGAGCTTGCAGTTGGCTCATTGGAATGAAATGCAATT 840
Db 800 ATTGAGCAGGAGTTTAAACAGAGCTTGCAGTTGGCTCATTGGAATGAAATGCAATT 859
QY 841 TTGATGAAACAGTACATTTAGTTTGTAGTACAGATTAATTTGGAATTTGGAATTTCA 900

860	TTGATGAAACAGTACATTGAGTTTGTAGCTGCAGAGTTACTTGTGGAACTTGTGAACTCTCA	919
Qy		
901	AAGGTTTTTTCAGGCAGAAAAATCCTTTTGATTTTATGAAAAACAATTTCTTTAGAGGAAAA	960
920	AAGGTTTTTTCAGGCAGAAAAATCCTTTTGATTTTATGAAAAACAATTTCTTTAGAGGAAAA	979
Qy		
961	ACAAATTTCTTTCAGAAAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC	1020
980	ACAAATTTCTTTCAGAAAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC	1039
Qy		
1021	ACAGATAACGCTCTTCACTTGGATGCAGATTTT	1053
1040	ACAGATAACGCTCTTCACTTGGATGCAGATTTT	1072
Db		

RESULT 7

```

US-09-325-301-505
; Sequence 505, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 505
; LENGTH: 1989
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1917)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-505

```

Query Match 54.9%; Score 578.4; DB 9; Length 1989;
Best Local Similarity 75.0%; Pred. No. 1e-156;
Matches 723; Conservative 0; Mismatches 241; Indels 0;

Qy	89	ATGAGAGCGCACTCCTTAAGAAAGAGTTCTTCGCGGGTTGTGCATCTTTCCCAATCCAGTACC	148
Db	280	AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTTCATCTTCCCATCGAGTACC	339
Qy	149	CTGATATTTGGAAAACTGATAAACAGACACAGGCTTCCTTCGACACAGCGAAGAGGTGG	208
Db	340	ATGATATCTGGCAGATGATAAAGAGGAGAGGCTTCCTTTTGACCGCCGAGGAGGTGG	399
Qy	209	ACTTATCAAAGGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAAGTACTTCATCT	268
Db	400	ACCTCTCAAGGACATTCAGCACATGGGATCCCTGAAACCCGAGGAGAGATATTTTATAT	459
Qy	269	CTCACATCTTAGCGCTTTTGTGGACCGATGATGAAATTGTAAATGAAAAATTTGGTGGAGC	328
Db	460	CCCATGTTCTGGCTTCTTTTGCAGCAACGATGGCATAGTAAATGAAAACTTGGTGGAGC	519
Qy	329	GCATTAGTCAGGAGGTCCAGGCTCCGAGGCTCGCTGTTCTATAGGCTTTCAAAATCTCA	388
Db	520	GATTTAGCCAAAGATTCAGATTAACAAGGCCGCTGTTCTATGGCTTTCAAATTTGCCA	579
Qy	389	TCGAGAAATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCA	448
Db	580	TGGAAACATACATTCCTGAAATGTATAGTCTCTTATTGGACACTTACATAAAGATCCCA	639
Qy	449	AGAAAGGGAAATTTTATATTTAAATGCAATTTGAAACCAATGCCCTATGTTAAGAAAAAGCAG	508
Db	640	AAGAAAGGGAAATTTCTCTTCAATGCCATTTGAAACGATGCCCTTGTGTCAAGAAAGAGGCAG	699

509	Qy	ATTGGGCGCTTGC	GATGGA	TAGCAGATAG	AAATCT	CTATTGGGGAAAG	TGCTGCGCCT	568
700	Db	ACTGGGCGCTTGC	CGCTGGA	TTGGGGCA	CAAGAGGCT	ACCTATGCTGAA	CGCTGTTGTAGCCT	759
569	Qy	TTGCTGCTG	TAGAAGAG	TTTTCTCT	CTCAGGATCT	TTTTGCTGCTAT	TATCTGCTGAAGA	628
760	Db	TTGCTGAG	TGGAGGCA	ATTTCT	TTTTCGGTTCT	TTTGGCTCGAT	TATCTGCTCAAGA	819
629	Qy	AGAGAGCTCT	TATGCCAGG	ACTCACT	TTTTTCCAATGAA	CTCATCAGCAGAGAT	GAAAGGAC	688
820	Db	AACGAGG	ACTATGCC	CTGGCCT	CACATTTTCT	TAATGAACTT	TATAGCAGAGATGAGGGTT	879
689	Qy	TTCACTGTG	ACTTTGCT	TGCTGAT	GTGTTCAAT	CTTAGTAATA	TAGACCTTCAGAGAAA	748
880	Db	TACACTGTG	ATTTTGTCT	TGCTGAT	GTTCAAACAC	CTGGTACACAA	ACCATCGGAGGAGA	939
749	Qy	GGGTGAGG	GAGATCAT	TTTGTG	TGCTCAAAAT	TGCAGCAGGAG	TGTTTTTAAACAGAGACCT	808
940	Db	GAGTAGAG	AAATAAT	CAATGCT	TTGTCGGATAG	AACAGGAGTT	CTCTCACTGAGGCCT	999
809	Qy	TGCCAGT	TGGCCTCAT	TGTGGAAT	GAAATGCA	TTTTTGATGAA	ACAGTACATTTGAGTTTGTAG	868
1000	Db	TGCCTGTG	AGCTCAT	TGTGGATGA	ATTGCACT	CTAATGAAGCA	ATACATTTGAGTTTGTGG	1059
869	Qy	CTGACAGAT	TACTTGTG	GAACTTGG	ATTTCTCA	AAAGTTTTT	CAGGCAGAGAAAATCCTTTTG	928
1060	Db	CAGACAG	ACTTATG	CTGGA	CTGGGT	TTTTAGCAAGG	TTTTTTCAGAGTAGAGAACCCATTTG	1119
929	Qy	ATTTTTTGA	AAAAAT	TTCTTTT	TAGAAGGAAAA	CAAAATTTCT	TTTTTGAGAACGAGTTTCAG	988
1120	Db	ACTTTAT	TGGAGAATA	TTTCT	CTGGAAGGAAG	ACTAACTTCT	TTTGTAGAGAGAGTAGGCG	1179
989	Qy	AGTATCAG	CGCTTTT	TGCACTTT	TATGGCGAAGAA	ACCA	CAGATAACGCTTTCACCTTGGATGCGAG	1048
1180	Db	AGTATCAG	AGGATGGG	AGTGATCT	CAAGTCC	AACACAGAGAA	TTCTTTTACCTTGGATGCTG	1239
1049	Qy	ATTT	1052					
1240	Db	ACTT	1243					

RESULT 8

```

RES001.8
US-10-084-817-342
; Sequence 342, Application US/10084817
; Publication No. US20030119009A1
GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
CURRENT APPLICATION NUMBER: US/10/084,817
CURRENT FILING DATE: 2002-02-25
PRIOR APPLICATION NUMBER: 60/270,784
PRIOR FILING DATE: 2001-02-23
NUMBER OF SEQ ID NOS: 365
SOFTWARE: PERL Program
SEQ ID NO 342
LENGTH: 2216
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 202239.1
US-10-084-817-342

```

Query Match	54.9%	Score 578.4;	DB 15;	Length 2216;
Best Local Similarity	75.0%;	Pred. No. 1.1e-156;		
Matches 723;	Conservative	0;	Mismatches 241;	Indels 0;
Gaps	0;			
Qy	89	ATGAAGAGGCACCTCCTAAGAAAGAGTTCTCGCGGTTTGTCACTTTTCCCAATCCAGTACC	148	

Db 284 AGGATGAGCCGCTGCTGAGAGAAACCCCGCGCTTTGTCTATCTTCCCATCGAGTACC 343
Qy 149 CTGATATTGGAAAATGTATAACAGGCACAGGCTTCTCTTGGACAGCAGAGAGGTGCG 208
Db 344 ATGATATCTGCGAGATGTATAAGAGGCGAGAGGCTTCTTTTGGACCCGCGAGAGGTGG 403
Qy 209 ACTTATCAAGGATCTCCCTCACTCGGAACAAGCTTAAAGCAGATGAGAAGTACTTCTCT 268
Db 404 ACCTCTCCAGGACATTCAGCACTGGGAATCCCTGAACCCGCGAGAGAGATATTTATAT 463
Qy 269 CTCACATCTTAGCCCTTTTTCAGCCAGTGTGAATTTGTAATGAAATTTGGTGGAGC 328
Db 464 CCCATGTTCTGGCTTTCTTTCAGCAAGCGATGATAGTAATGAAATTTGGTGGAGC 523
Qy 329 GCTTTAGTCAGGAGGTGCGAGTTCAGAGGCTGCTGTTCTATGGCTTTTCAAAATCTCA 388
Db 524 GATTTAGCCAGAGTTTCAGATTACAGAGCCGCTGTTCTATGGCTTCCAAATTTGCCA 583
Qy 389 TCGAATGTTTCACTCAGAGATGTACAGTGTGCTGATAGACACTTACATCAGAGATCCCA 448
Db 584 TGGAAAACATACATTCGAATGTATAGTCTTCTATTGACACTTACATAAAGATCCCA 643
Qy 449 AGAAAGGGAAATTTTATTAATGCAATTTGAACCATGCTCTATGTTTAAAGAAAAGCAG 508
Db 644 AAGAAAGGGAAATTTCTCTTCAATGCCATTTGAACGATGCTTGTGTCAAGAAAGAGCGAG 703
Qy 509 ATTGGGCTTGGCATGATAGATAGATAGAAATCTACTTTTGGGGAAGAGTGGTGCT 568
Db 704 ACTGGGCTTGGCTGATTTGGGACAAAGAGGCTACCTATGTTGAACGTTGTAGGCT 763
Qy 569 TTGCTGCTGTAGAGAGTGTCTTCTCAGGATCTTTTGTGCTGATATTTCTGGCTTAAAGA 628
Db 764 TTGCTGAGTGGAGGCAATTTCTTTTCGGTCTTCTTTGGCTGATATTTCTGGCTCAAGA 823
Qy 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGAC 688
Db 824 AACGAGGACTGATGCTGGCTGCATTTTCTAATGAACCTTATTAGCAGAGATGAGGTT 883
Qy 689 TTCACTGTGACTTTGCTTGGCTGATGTTCCAAATGAACTTATTAATGAAGCTTCAAGAAA 748
Db 884 TACACTGTGATTTTGGCTTGGCTGATGTTCAAAACACCTGGTACACAAACCATCGGAGAGA 943
Qy 749 GGGTCAGGAGATCAATTTGTGATGCTCTCAAAATTTGAGCAGGAGTGTTTTAAACAGAGCT 808
Db 944 GAGTAAGAGAAATTAATTAATCAATGCTGTGCGATAGAACAGAGTTCCTCACTGAGGCT 1003
Qy 809 TGCAGTGGGCTCAATTTGGAATGAATTCATTTTGAATGAACAGTACATGAGTGTGTAG 868
Db 1004 TGCCTGTGAAGCTCAATTTGGGATGAATTTGCACTCTAATGAAGCAATACATGAGTGTGTGG 1063
Qy 869 CTGACAGATTACTTGTGGAATCTGATTTCTCAAGGTTTTCAGGCGAGAAATCTTTTG 928
Db 1064 CAGACAGACTTATGCTGGAATCTGGGTTTATGCAAGGTTTTCAGAGTAGAGAACCCATTG 1123
Qy 929 ATTTTATGGAACATTTCTTTAGAGGAAAACAAATTTCTTTTGAAGAAACGAGTTTCAG 988
Db 1124 ACTTTATGAGATATTTTCTGAGAGAAAGACTAATCTTTTGAAGAGAGTAGCGG 1183
Qy 989 AGTATCAGCGTTTTCAGTGTATGSCAGAAACCAAGATAAGCTTTCACCTTGGATGCGAG 1048
Db 1184 AGTATCAGAGGATGGGAGTGTATGATCAAGTCCAACAGAGAAATTTCTTTTACCTTGGATGCTG 1243
Qy 1049 ATTT 1052
Db 1244 ACTT 1247

RESULT 9

US-10-335-514/c

; Sequence 514, Application US/10220335

; Publication No. US20050175607A1

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides

; FILE REFERENCE: 21272-041

; CURRENT APPLICATION NUMBER: US/10/220,335

; CURRENT FILING DATE: 2002-08-28

; PRIOR APPLICATION NUMBER: 09/664,641

; PRIOR FILING DATE: 2000-09-19

; PRIOR APPLICATION NUMBER: 09/616,807

; PRIOR FILING DATE: 2000-07-14

; PRIOR APPLICATION NUMBER: 09/597,707

; PRIOR FILING DATE: 2000-06-17

; PRIOR APPLICATION NUMBER: 09/577,409

; PRIOR FILING DATE: 2000-05-18

; PRIOR APPLICATION NUMBER: 09/515,126

; PRIOR FILING DATE: 2000-02-28

; NUMBER OF SEQ ID NOS: 688

; SOFTWARE: Custom

; SEQ ID NO 514

; LENGTH: 2482

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-10-220-335-514

Query Match

Best Local Similarity 54.9%; Score 578.4; DB 22; Length 2482;

Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

Qy 89 ATGAGAGCCACTCTCTAAGAAAGAGTCTCTGCGCGCTTGTCTCATCTTTCCAAATCCAGTACC 148
Db 2086 AGGATGAGCCGCTGCTGAGAGAAAACCCCGCGCTTGTCTCATCTTTCCCATCGAGTACC 2027
Qy 149 CTGATATTGGAAAATGTATAACAGGCACAGGCTTCTTCTGGACAGCAGAGAGGTGCG 208
Db 2026 ATGATATCTGCGAGATGTATAAGAGGCGAGGCTTCTTTTGGACCGCGAGAGGTG 1967
Qy 209 ACTTATCAAGGATCTCCCTCACTGGAACAAGCTTTAAAGCAGATGAGAAGTACTTCTCATCT 268
Db 1966 ACCTCTCAAGGACATTCAGCACCTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 1907
Qy 269 CTGACATCTTAGCCCTTTTTCAGCCAGTGTGAATTTGTAATGAAAATTTGGTGGAGC 328
Db 1906 CCCATGTTCTGGCTTTCTTTTGCAGCAAGCATGGCATAGTAATAATGAAAATTTGGTGGAGC 1847
Qy 329 GCTTTAGTCAGAGGTGCGAGTTCAGAGGCTCGCTGTTTCTATGCTTTCATAAATCTCA 388
Db 1846 GATTTAGCCAAAGAGTTCAGATTCAGAAAGCCGCTGTTTCTATGGCTTCCAAATTTGCCA 1787
Qy 389 TCGAAGATGTTTCACTCAGAGATGTACAGTGTGCTGATAGACACTTACATCAGAGATCCCA 448
Db 1786 TGGAAAACATACATTCGAAATGTATAGTCTTCTTATTGACACTTACATAAAGATCCCA 1727
Qy 449 AGAAAAGGGAATTTTATTATTAATGCAATTTGAAACCAATGCCCTATGTTTAAAGAAAAGCAG 508
Db 1726 AAGAAAGGGAATTTCTCTTCAATGCCAATTTGAAACGATGCTTGTGTCAAGAAAGAGCAG 1667
Qy 509 ATTGGGCTTCCGATGATAGCAGATGAGAAATCTACTTTTGGGGAAGAGGTGGTGGCCT 568
Db 1666 ACTGGGCTTCCGCTGGATTTGGGACAAAGAGGTACCTATGTTGTAAGAGTGTGTAGCCT 1607
Qy 569 TTGCTGCTGTGAAGAGGAGTTTTCTTCTCAGGATCTTTTGTGCTGCTATATTTCTGCTTAAAGA 628
Db 1606 TTGCTGAGTGGAGGCAATTTCTTTTCCGGTCTTTTGGCTGATATTTCTGGCTCAAGA 1547
Qy 629 AGAGAGTCTTATGCGCAGGACTCACTTTTTCATGAACTCATCAGCAGAGATGAGAGC 688
Db 1546 AACGAGGACTGATGCTGGCTCTACATTTTCTAATGAACCTTATTAGCAGAGATGAGGTT 1487
Qy 689 TTCACTGTGACTTTGCTTGGCTGATGTTTCCAAATCTTAGTAATTAAGCCTTCAGAGAAA 748
Db 1486 TACACTGTGATTTTGGCTTGGCTGATGTTTCAAAACCTGTGTACAAAACCATCGAGAGAGA 1427
Qy 749 GGGTCAGGAGATCAATTTGTTGATGCTGTCAAAATTTGAGCAGGAGTTTTAAACAGAGCCT 808
Db 1426 GAGTAAGAAATTAATTTATCAATGCTGTTCCGATAGAACAGGAGTTCCTCACTGAGGCT 1367

;; PRIOR FILING DATE: 2000-09-18
;; PRIOR APPLICATION NUMBER: US/60/234,052
;; PRIOR FILING DATE: 2000-09-20
;; PRIOR APPLICATION NUMBER: US/60/234,923
;; PRIOR FILING DATE: 2000-09-25
;; PRIOR APPLICATION NUMBER: US/60/235,134
;; PRIOR FILING DATE: 2000-09-25
;; PRIOR APPLICATION NUMBER: US/60/235,637
;; PRIOR FILING DATE: 2000-09-26
;; PRIOR APPLICATION NUMBER: US/60/235,638
;; PRIOR FILING DATE: 2000-09-26
;; PRIOR APPLICATION NUMBER: US/60/235,711
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: US/60/235,720
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: US/60/235,840
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: US/60/235,863
;; PRIOR FILING DATE: 2000-09-27
;; NUMBER OF SEQ ID NOS: 2276
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 1169
;; LENGTH: 2500
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-954-456-1169

Query Match 54.9%; Score 578.4; DB 9; Length 2500;

Best Local Similarity 75.0%; Pred. No. 1.1e-156; Mismatches 241; Indels 0; Gaps 0;

89 ATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGCCGGTTGTCTCATCTTTCCAAATCCAGTACC 148
397 AGGATGAGCGCTCTCTAGAGAAACCCCGCCCTTGTCTCATCTTTCCCAATCCAGTACC 456
149 CTGATATTTGGAAATGTATAACAGGCACAGGCTTCTCTCGACAGCAGAGAGGTG 208
457 ATGATATCTGCAGATGTATAAGAGGCAGAGGCTTCTTTTGGACCCGAGGAGGTG 516
209 ACTTATCAAGGATCTCCCTCACTGGAACAGCTTAAAGCAGATGAGAGTACTTCACT 268
517 ACCTCTCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGCAGAGATATTTATAT 576
269 CTCACATCTTAGCCCTTTTTCGACGAGTGAATTTGTAATGAAATTTGCTGAGC 328
577 CCCATGTTCTGGCTTCTTTTCGACGAGCGATGCAATAGTAATGAAACCTTGTGAGC 636
329 GCTTTAGTCAGGAGTGCAGGTTCCAGAGGCTCGCTGTCTTATGGCTTTCCAAATTTCTCA 388
637 GATTTAGCAAGAGTTCAAGTACAGAGCCGCTGTCTTATGGCTTCCAAATTTGCCA 696
389 TCGAAGATGTTCACTCAGAGATGACAGTTTGTCTGATAGACACTTACATCAGAGATCCCA 448
697 TGGAAAACATACATCTCTGAAATGTATAGTCTCTTATTTAGACACTTACATAAAGATCCCA 756
449 AGAAAGGGATTTTATTAATGCAATTTGAACCATGCCCTATGTTAAGAAAAGCAG 508
757 AAGAAAGGGAATTTCTCTCAATGCCATTTGAACGATGCCCTTGTGTCAAGAAAGAGCAG 816
509 ATTGGGCTTTCGATGATAGCAGATAGAAATCTACTTTTGGGGAGAGAGTGTGGCCT 568
817 ACTGGGCTTTCGCTGATTTGGGACAAAGAGGCTACTATGGTGAACGTGTTGAGCCT 876
569 TTGCTGCTGTAGAGGAGTTTCTTCTCAGGATCTTTTGTGCTATATTTCTGGCTAAAGA 628
877 TTGCTGAGTGAAGGCAATTTCTTTTTCGGTCTTTTGGTCTGATATTTCTGGCTCAGA 936
629 AGAGAGTCTTATGCCAGGACTCACTTTTCCAAATGAATCTCATCAGCAGAGATGAAGAC 688
937 AACGAGGACTGATGCTGGCTCACAATTTTCTAATGAATTTATATAGCAGATGAGGTT 996
689 TTCACTGAGACTTTGCTTGTGCTGATGTTCCAAATCTTAGTAATTAAGCCTTCAAGAA 748

Db 997 TACACTGTGATTTTCTGCTTGCCTGTGATGTTCAAACACCTGTGACACAAACCAATCGAGGAGA 1056
Qy 749 GGGTCAGGAGATCAATTTGTTGATGCTGTCAAAATTTGAGCAGAGTGTCTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATAATTAATCAATGCTGTTCGATAGACAGAGTTCCTCACTGAGCCT 1116
Qy 809 TGCCAGTTGGCCTCAATTTGGAATGAATTTGCAATTTTGAATGAAACAGTACATTTGTTAG 868
Db 1117 TGCCTGTGAAGCTCAATTTGGATGAATTTGCACTCTAATGAAGCAATACATTTGAGTTGTGG 1176
Qy 869 CTGACAGATTTACTTTGTTGGAATTTGGAATTTCTCAAGGTTTTCAGGCAGAAAATCCTTTT 928
Db 1177 CAGACAGACTTTATGCTGGAATTTTGTAGCAAGGTTTTCAGAGTAGAGAACCCATTTG 1236
Qy 929 ATTTTATGAAAACATTTCTTTAGAGGAAACAAAATTTCTTTTGAAGAACAGAGTTTCAG 988
Db 1237 ACTTTATGAGATATTTTCACTGGAGGAAAGACATACTTCTTTGAGAGAGAGTAGGCG 1296
Qy 989 AGTATCAGCGTTTTCAGTATTTGCGAGTATTTGCGAGAAACACAGATAACGCTTTTACCTTGGATGCAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGTCAAGTCCCAACAGAGAAATTTCTTTTACCTTGGATGCAG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 12

US-09-954-456-1827
; Sequence 1827, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1827
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1827

Query Match 54.9%; Score 578.4; DB 9; Length 2500;

Best Local Similarity 75.0%; Pred. No. 1.1e-156; Mismatches 241; Indels 0; Gaps 0;

89 ATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGCCGGTTGTCTCATCTTTCCAAATCCAGTACC 148
397 AGGATGAGCGCTCTCTAGAGAAACCCCGCCCTTGTCTCATCTTTCCCAATCCAGTACC 456

QY 509 ATTGGGCTTGGCGATGATAGCAGATAGAAATCTACTTTTGGGAAAGAGTGGGCT 568
DB 817 ACTGGGCTTGGCGATGATGGGACAAAGAGGCTACCTATGGTGAAGTGTAGGCT 876
QY 569 TTGCTGCTGTAGAAGGATTTCTTCTCAGAGCTTTTGGCTATATCTTGGCTAAAGA 628
DB 877 TTGCTGAGTGAAGGATTTCTTCTTGGCTATATCTTGGCTCAAGA 936
QY 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCAAATGAATCTCATCAGCAGAGATGAAGGAC 688
DB 937 AACGAGAGCTGATGGCTGGCTCACAATTTCTAATGAATCTTATAGCAGATGAGGTT 996
QY 689 TTCACTGTGACTTTTGGCTTGGCTGATGTTCCAAATCTTAGTAAATAAGCCTTCAGAAGAAA 748
DB 997 TACACTGTGATTTTGGCTTGGCTGATGTTCCAAACACCTGGTACACAAACCATCGGAGGA 1056
QY 749 GGGTCAGGAGATCATTTGCTGCTGATGCTGCTCAAAATGTAGAGCAGGATTTTAAACAGAGCCT 808
DB 1057 GAGTAAGAGAAATAATATCAATGCTGTTGGATAGAACAGGAGTTCTCACTGAGGCT 1116
QY 809 TGCCAGTTGGCCTCATTTGGGATGAATGCTATTTGATGAACACATACATTGAGTTGTAG 868
DB 1117 TGCTGTGAAGCTCATTTGGGATGAATGCTATTTGATGAACACATACATTGAGTTGTAG 1176
QY 869 CTGACAGATTAATGCTGGAATCTTCAAGGTTTTCAGGAGAAATCCTTTTG 928
DB 1177 CAGACAGATTAATGCTGGAATCTTCAAGGTTTTCAGGAGAAATCCTTTTG 1236
QY 929 ATTTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
DB 1237 ACTTTATGGAATATTTCTCACTGGAAGGAAAGCTAACTTTCTTTGAGAGAGTAGGCG 1296
QY 989 AGTATCAGGCTTTGCGAGTTATGCGAGAACCAACAGATACGTTTCACTTGGATGAG 1048
DB 1297 AGTATCAGGAGTGGAGTGTGATGTCAGTCCAAACAGAGAAATTTCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
DB 1357 ACTT 1360

RESULT 14

US-10-733-878-458
; Sequence 458, Application US/10733878
; Publication No. US20040224408A1
; GENERAL INFORMATION:
; APPLICANT: Jean-Philippe Girard
; APPLICANT: Francois Amalric
; APPLICANT: Myriam Rousseau
; APPLICANT: Thomas Clouaire
; TITLE OF INVENTION: THAP PROTEINS AS NUCLEAR RECEPTORS FOR
; TITLE OF INVENTION: CHEMOKINES AND ROLES IN TRANSCRIPTIONAL REGULATION, CELL
; TITLE OF INVENTION: PROLIFERATION AND CELL DIFFERENTIATION
; FILE REFERENCE: BIOBANK.012A
; CURRENT APPLICATION NUMBER: US/10733,878
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 60/432699
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/485027
; PRIOR FILING DATE: 2003-07-03
; NUMBER OF SEQ ID NOS: 535
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 458
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-733-878-458

Query Match 54.9%; Score 578.4; DB 20; Length 2500;
Best Local Similarity 75.0%; Pred. No. 1.1e-156;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCCACTCTCTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTTTCCAAATCCAGTACC 148
DB 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTCTATCTTTCCCATCGAGTACC 456
QY 149 CTGATATTTGGAAATGATATAAAGAGGACAGGCTTCTTCTGGAACAGCAGAGAGAGTGC 208
DB 457 ATGATATCTGGCAGATGATATAAAGAGGACAGGCTTCTTCTTGGACCGCGAGAGGTTG 516
QY 209 ACTTATCAAAAGGATCTCCCTCACTGGAACAAGCTTAAACAGATGAGAGATCTTCACTCT 268
DB 517 ACCTCTCAAGAGCAATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
QY 269 CTCATCTTAGCCTTTTTCAGGCCAGTGAATTTGTAATGAAATTTTGGTGGAGC 328
DB 577 CCCATGTTCTGGCTTTCTTTGAGCAAGCATAGGATAGTAAATGAAACTTGGTGGAGC 636
QY 329 GCTTTAGTCAGAGGTGCGAGGCTTCCAGAGGCTCGCTTCTTATGAGCTTTCAAAATCTCA 388
DB 637 GATTTAGCCAAAGAGTTTCAGATTAACAGAGCCCGCTGTTTCTATGAGCTTCCAAATTGCCA 696
QY 389 TCGAAGATGTTCACTCAGAGATGATGCTTGTGATGACACTTACATCAGAGATCCCA 448
DB 697 TGGAAACATACATCTGAAATGTATAGTCTTTTATGACACTTACATATAAAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTTAATGCAATTTGAAACCAATGCCCTATGTTAAAGAAAGAGCAG 508
DB 757 AAGAAAGGGAATTTCTTCAATGCCATTTGAAACGATGCTTGTCTCAAGAGAGAGCAG 816
QY 509 ATTGGGCTTCCGATGATAGCAGATAGAAATCTACTTTTGGGAAAGAGTGGTGGCCT 568
DB 817 ACTGGGCTTGGCGTGGATGGGACAAAGAGGCTACCTATGTTGAACGCTGTTGTAGCCT 876
QY 569 TTGCTGCTGTAGAGGAGTTTCTTCTCAGAGATCTTTTGTGCTGCTATATTTCTGGCTAAAGA 628
DB 877 TTGCTGAGTGGAGGCAATTTCTTTTCCGGTTCTTTTGGCTGCTATTTCTGGCTCAAGA 936
QY 629 AGAGAGTCTTATGCGAGGCTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGGAC 688
DB 937 AACGAGAGCTGATGCTGGCTCACTTTTCTAATGAATCTTATAGCAGAGATGAGGTT 996
QY 689 TTCACTGTGACTTTTGGCTTGGATCTTCAATACCTTAGTAAATAAGCCTTCAGAAGAAA 748
DB 997 TACACTGTGATTTTGGCTTGGATCTTCAACACCTGCTACACAAACCATCGGAGGAGA 1056
QY 749 GGGTCAGGAGATCATTTGCTGATGCTGCAAAATTTGAGCAGAGGTTTAAACAGAGCCT 808
DB 1057 GAGTAAGAGAAATAATATCAATGCTGTTGGATAGAACAGGAGTTCTCTCACTGAGGCT 1116
QY 809 TGCCAGTTGGCCTCATTTGGGATGAATGCTATTTGATGAACACATACATTGAGTTGTAG 868
DB 1117 TGCTGTGAAGCTCATTTGGGATGAATGCTATTTGATGAACACATACATTGAGTTGTAG 1176
QY 869 CTGACAGATTAATGCTGGAATCTTCAAGGTTTTCAGGAGAAATCCTTTTG 928
DB 1177 CAGACAGATTAATGCTGGAATCTTCAAGGTTTTCAGGAGAAATCCTTTTG 1236
QY 929 ATTTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
DB 1237 ACTTTATGGAATATTTCTCACTGGAAGGAAAGCTAACTTTCTTTGAGAGAGTAGGCG 1296
QY 989 AGTATCAGGCTTTGCGAGTTATGCGAGAACCAACAGATACGTTTCACTTGGATGAG 1048
DB 1297 AGTATCAGAGGATGGAGTGTGATGTCAGTCCAAACAGAGAAATTTCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
DB 1357 ACTT 1360

RESULT 15
US-10-843-641A-3751
; Sequence 3751, Application US/10843641A
; Publication No. US20050064454A1

GENERAL INFORMATION:

APPLICANT: Avalon Pharmaceuticals, Inc.
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
TITLE OF INVENTION: Signature Gene Sets
FILE REFERENCE: 689290-189
CURRENT APPLICATION NUMBER: US/10/843,641A
CURRENT FILING DATE: 2004-05-12
PRIOR APPLICATION NUMBER: US/09/873,367
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US/09/954,531
PRIOR FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/09/954,456
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US/09/962,436
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US/09/962,832
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US/09/964,824
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: US/09/967,768
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US/09/968,007
PRIOR FILING DATE: 2001-10-02
PRIOR APPLICATION NUMBER: US/09/969,347
PRIOR FILING DATE: 2001-10-02
PRIOR APPLICATION NUMBER: US/09/969,708
PRIOR FILING DATE: 2001-10-03
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 8447
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3751
LENGTH: 2500
TYPE: DNA
ORGANISM: Homo sapiens
US-10-843-641A-3751

Query Match 54.9%; Score 578.4; DB 21; Length 2500;
Best Local Similarity 75.08; Pred. No. 1.1e-156;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCACTCTAAGAAAGAGTTCTCGCGGTTTGTCACTTTTCCAATCCAGTACC 148
DB |||||
QY 149 CTGATATTGGAAATGTATAACAGGCACAGGCTTCTCTGACACAGCAGAGAGTCG 208
DB |||||
QY 457 ATGATATCTGGCAGATGATAAGAAAGCAGAGGCTTCTTTTGGACCGCGGAGGTTG 516
DB |||||
QY 209 ACTTATCAAAGGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTACTTCATCT 268
DB |||||
QY 517 ACCTCTCAAGACATTCAGCACTGGGAAATCCCTGAAACCGAGGAGAGATATTTTATAT 576
DB |||||
QY 269 CTCACATCTTAGCCTTTTTCAGCCAGTGTATGGAATTTGTAATGAAATTTGGTGAGC 328
DB |||||
QY 577 CCCATGTTCTGGCTTTCTTTGACGACGATGCGATGAGTAAATGAAATCTTGGTGAGC 636
DB |||||
QY 329 GCTTTAGTACAGAGTCAGGTTCCAGAGGCTCGCTGTTCTATGGCTTTCAAAATCTCA 388
DB |||||
QY 637 GATTTAGCCAAAGAGTTTCAATACAGAACGCCGCTGTTTATGGCTTCCAAATTTGCCA 696
DB |||||
QY 389 TCGAGATGTTCTACAGAGTGTACAGTTTGTGTATAGACACCTTACATCAGAGATCCCA 448
DB |||||
QY 697 TGAATAACATACATCTGAAATGTATAGTCTTCTTATGTACACTTACATAAAGATCCCA 756
DB |||||
QY 449 AGAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCGCTATGTTAAGAAAGAGCAG 508
DB |||||
QY 757 AAGAAAGGGAATTTCTCTTCAATGCCATTTGAACGATGCCCTTGTCTCAAGAGAGCAG 816
DB |||||
QY 509 ATTGGGCTTGGAGTACAGATAGAAATCTATTTTGGGAAAGAGTGGTGGCCT 568
DB |||||
QY 817 ACTGGGCTTGGGCTGAGTTGGGACAAAGAGGCTACCTATGGTGAACGTTGTAGCCT 876
DB |||||
QY 569 TTGCTGCTGTAGAAGAGTTTCTTCTCAGGATCTTTTGTGCTATATTTCTGGCTAAGA 628
DB |||||

DB 877 TTGCTGCAGTGAAGGCAATTTTCTTTTCCGGTCTTTTTCGGTTCGATATTTCTGCTCAAGA 936
QY 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCAAATGAATCACTCATCAGCAGAGATGAAGGAC 688
DB 937 AACGAGGACTGATGCTGGCCTCACATTTTCTAATGAATTTATTAGCAGAGATGAGGGTT 996
QY 689 TTCACTGTGACTTTGCTTGGCTGATGTTCCCAATCTTAGTAAATAAGCCTTCAGAAGAAA 748
DB 997 TACACTGTGATTTTGGTTCCTGATGTTCAAAACACCTGCTGTACAAACCATCGAGGAGA 1056
QY 749 GGGTCAGGAGATCAATTTGATGCTGTCAAAATTTGAGCAGGAGTTTAAACAGAGCCT 808
DB 1057 GAGTAAGAGAAATTAATTAATCAATGCTGTTCGGATAGAACAGGAGTTCTCTCAGGCT 1116
QY 809 TGCAGTTGGCCTCAATTTGGAATGAATTTGCATTTTGAATGAATAAGTACATTTGATTTGTAG 868
DB 1117 TGCCTGTGAAGCTCAATTTGGGATGAATTTGCACTCTAATGAAGCAATACATTTGATTTGTGG 1176
QY 869 CTGACAGATTACTTTGTGGAATCTTCAAGGTTTTCAGGCAGAAATCCCTTTG 928
DB 1177 CAGACAGACTTATGCTGGAATCTGGGTTTTCAGAGGTTTTCAGAGTAGAGAACCCATTTG 1236
QY 929 ATTTTATGGAACAACTTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTTCAG 988
DB 1237 ACTTTATGGAGATATTTCTCTGGAAGGAAAGACTAATCTTTTGAAGAGAGTAGGCG 1296
QY 989 AGTATCAGCGTTTTCAGTTATGAGGAGAAACACAGATAGCTTTCACCTTTGGATGACAG 1048
DB 1297 AGTATCAGAGGATGGGAGTGTGTCAAGTCCAACAGAGAAATTTCTTTTACCTTTGGATGCTG 1356
QY 1049 ATTT 1052
DB 1357 ACTT 1360

Search completed: September 26, 2005, 03:50:46
Job time : 583 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2005, 23:52:05 ; Search time 2732 Seconds
(without alignments)
14671.184 Million cell updates/sec

Title: US-10-698-228-2
Perfect score: 1053
Sequence: 1 atggcgaccggaagcc.....tcacctggatgcagatttt 1053

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gsl1.*
9: gb_gsl2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1053	100.0	3292	3	CR617553 full-leng
2	1053	100.0	4748	3	AL137348 Homo sapi
3	1041	98.9	4650	3	BC042948 Homo sapi
4	1010.6	96.0	3501	3	CR627376 Homo sapi
5	945	89.7	1056	9	AY398973 Homo sapi
6	929.6	88.3	1083	1	AL547501
7	829.8	78.8	1008	9	AY398974
8	793.2	75.3	879	5	BC041857
9	777	73.8	1043	9	AY398975
10	731.2	69.4	892	7	CN163214
11	728.2	69.2	914	7	CF995079
12	710.6	67.5	870	7	CO648380
13	682.8	64.8	804	6	CD656661
14	655.4	62.2	1061	4	EM801298
15	654.8	62.2	896	5	BU170979
16	616.4	58.5	685	5	BX951374
17	578.4	54.9	1571	3	CR603461
18	578.4	54.9	1573	3	CR625489
19	578.4	54.9	1582	3	CR608076
20	578.4	54.9	1588	3	CR602054
21	578.4	54.9	1592	3	CR621427
22	578.4	54.9	1600	3	CR604378
23	578.4	54.9	1605	3	CR590959
24	578.4	54.9	1612	3	CR614990

25	578.4	54.9	1613	3	CR609838	CR609838 full-leng
26	578.4	54.9	1623	3	CR618451	CR618451 full-leng
27	578.4	54.9	1630	3	CR602150	CR602150 full-leng
28	578.4	54.9	1796	3	CR596700	CR596700 full-leng
29	570.4	54.2	2088	3	AK088907	AK088907 Mus muscu
30	562	53.4	638	7	CN358022	CN358022 170005831
31	547	51.9	547	7	CN358025	CN358025 170005326
32	530.8	50.4	836	7	CR416843	CR416843 CR416843
33	528.8	50.2	608	1	AU137050	AU137050 AU137050
34	527.2	50.1	540	5	BU429508	BU429508 UI-HF-BNO
35	525	49.9	997	4	BM468712	BM468712 AGENCOURT
36	518.8	49.3	522	5	BQ694828	BQ694828 1000971 H
37	513.6	48.8	793	7	CK778582	CK778582 965745 MA
38	503.8	47.8	923	5	CR580780	CR580780 CR580780
39	503	47.8	792	5	BU328458	BU328458 603494277
40	502	47.7	510	7	CR540284	CR540284 DKFZp459N
41	501.4	47.6	697	6	CA328968	CA328968 UI-M-FY0
42	493.8	46.9	510	5	BX955728	BX955728 DKFZp781G
43	493.2	46.8	946	7	CN024619	CN024619 AGENCOURT
44	491.8	46.7	932	5	BU196941	BU196941 AGENCOURT
45	489.4	46.5	909	6	CA981614	CA981614 AGENCOURT

ALIGNMENTS

RESULT 1
CR617553
LOCUS full-length cDNA clone CSODI011YF14 of Placenta Cot 25-normalized
DEFINITION of Homo sapiens (human).
ACCESSION CR617553
VERSION CR617553.1 GI:50498360
KEYWORDS HTC; CNSLT_CDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (Bases 1 to 3292)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2. (Bases 1 to 3292)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
Location/Qualifiers
1. .3292
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI011YF14"
/tissue type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

REFERENCE 100.0%; Score 1053; DB 3; Length 3292;
Best Local Similarity 100.0%; Pred. No. 3.3e-269;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGGCGACCGGAAAGCGCGGCGGCTGGATCAGGATGAGATCATCTTCA 60
|||||
Db 27 ATGGGCGACCGGAAAGCGCGGCGGCTGGATCAGGATGAGATCATCTTCA 86

Query Match 100.0%; Score 1053; DB 3; Length 3292;
Best Local Similarity 100.0%; Pred. No. 3.3e-269;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGGCGACCGGAAAGCGCGGCGGCTGGATCAGGATGAGATCATCTTCA 60
|||||
Db 27 ATGGGCGACCGGAAAGCGCGGCGGCTGGATCAGGATGAGATCATCTTCA 86

Qy	61	GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC	120
Db	87	GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC	146
Qy	121	CGGTTTGTCTATCTTTCCAAATCCAGTACCCCTCATATTTGGAAATATGTAATAACAGGCACAG	180
Db	147	CGGTTTGTCTATCTTTCCAAATCCAGTACCCCTCATATTTGGAAATATGTAATAACAGGCACAG	206
Qy	181	GCTTCCTTTCTGGACAGCAGAGAGGTCCACTTATCAAAGGATCTCCCTCACTGCAACCAAG	240
Db	207	GCTTCCTTTCTGGACAGCAGAGAGGTCCACTTATCAAAGGATCTCCCTCACTGCAACCAAG	266
Qy	241	CTTAAGCAGATGAGAAGTACTTCAATCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT	300
Db	267	CTTAAGCAGATGAGAAGTACTTCAATCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT	326
Qy	301	GGAATGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCCAGAGGCT	360
Db	327	GGAATGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCCAGAGGCT	386
Qy	361	CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG	420
Db	387	CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG	446
Qy	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAAATTTTATTTAATGCAATTGAA	480
Db	447	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAAATTTTATTTAATGCAATTGAA	506
Qy	481	ACCATGCCCTATGTTAAGAAAAAGCAGATGGCCCTTGCAGATGATAGAGATAGAAAA	540
Db	507	ACCATGCCCTATGTTAAGAAAAAGCAGATGGCCCTTGCAGATGATAGAGATAGAAAA	566
Qy	541	TCTACTTTTGGGAAAGAGTGGTGGCCCTTGTCTGCTGTAGAGAGTTTCTTCTCAGGA	600
Db	567	TCTACTTTTGGGAAAGAGTGGTGGCCCTTGTCTGCTGTAGAGAGTTTCTTCTCAGGA	626
Qy	601	TCTTTTGTCTATATTCTGGCTAAAGAAAGAGAGGTCTTATGCCAGGACTCACTTTTTCC	660
Db	627	TCTTTTGTCTATATTCTGGCTAAAGAAAGAGAGGTCTTATGCCAGGACTCACTTTTTCC	686
Qy	661	AATGAACCTATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTTCCAA	720
Db	687	AATGAACCTATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTTCCAA	746
Qy	721	TACTTAGTAATAGCCTTCAGAAAGAGGTCAAGGAGATCATTTGTGATGCTGCTCAA	780
Db	747	TACTTAGTAATAGCCTTCAGAAAGAGGTCAAGGAGATCATTTGTGATGCTGCTCAA	806
Qy	781	ATTGAGCAGGAGTTTAAACAGAGCCCTTGCAGTGGCCCTCATTTGGAATGCAATTGCA	840
Db	807	ATTGAGCAGGAGTTTAAACAGAGCCCTTGCAGTGGCCCTCATTTGGAATGCAATTGCA	866
Qy	841	TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGTGGAACCTTGGATTCTCA	900
Db	867	TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGTGGAACCTTGGATTCTCA	926
Qy	901	AAGGTTTTTTCAGGCAGAAATCCTTTTGAATTTATGGAATAACATTTCTTTAGAGGAAA	960
Db	927	AAGGTTTTTTCAGGCAGAAATCCTTTTGAATTTATGGAATAACATTTCTTTAGAGGAAA	986
Qy	961	ACAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTATGCGCAGAAACC	1020
Db	987	ACAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTATGCGCAGAAACC	1046
Qy	1021	ACAGATAACGCTTTTCACCTTGGATGCAGATTTT	1053
Db	1047	ACAGATAACGCTTTTCACCTTGGATGCAGATTTT	1079
RESULT 2			
HSMB02033			
LOCUS			
HSMB02033		4748 bp	linear
		mRNA	HTC 22-SEP-2004

RESULT 2
HSMB02033

DEFINITION	Homo sapiens mRNA; cDNA DKFZp761E1312 (from clone DKFZp761E1312).
ACCESSION	AL137348
VERSION	AL137348.1 GI:6807859
KEYWORDS	HTC.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 4748) Ansorge, W., Krieger, S., Regiert, T., Rittmueller, C., Schwager, B., Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.
CONSTRM	The German cDNA Consortium
TITLE	Direct Submission
JOURNAL	Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp761E1312) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp761E1312 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.
FEATURES	Location/Qualifiers
source	1..4748 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="RZPD:DKFZp761E1312" /db_xref="taxon:9606" /clone="DKFZp761E1312" /tissue_type="amygdala" /clone_lib="761 (synonym: hamy2). Vector pSport1; host DH10B; sites NotI + SalI" /dev_stage="adult" /note="ribonucleotide reductase"
gene	1..4748 /gene="DKFZp761E1312"
CDS	47..1102 /gene="DKFZp761E1312" /codon_start=1 /product="hypothetical protein" /protein_id="CAB70703.2" /db_xref="GI:52545612" /db_xref="GOA:Q9NTD8" /db_xref="UniProt/TrEMBL:Q9NTD8" /translation="MGDPERPEAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFV PPIQYDIDWKYKQAOASFTWAEVDLSKDLPHWNKLKADKYPFISHILAFPAASDGI VNEILVERFSQEVQPERCYFGVQILLIENHSEMYSLIDITYIDPKRREPLFNALIE TMFYVKKADWALRWIADRKSTYGERVAFAPAAVEGFFSGSFAAIFWLKRGRLMPLGIT FSNELISRDEGLHCDFAELQFYLWKNPSEERVRELIIVDAVKIEQSLFTLEALPVGLIG MNCILMKQYIEFVADRLILVELGFSKVFQAEFPDFMENISLEGKTNFPEKRVSEYQRF AVMAETTDNVFTLDAPF"
ORIGIN	
Query Match	100.0%; Score 1053; DB 3; Length 4748;
Best Local Similarity	100.0%; Pred. No. 3.7e-269;
Matches 1053; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 ATGGGCGACCCGGAAGCCCGGAAGCCCGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Db	47 ATGGGCGACCCGGAAGCCCGGAAGCCCGCGGCTGGATCAGATGAGAGATCATCTTCA 106
Qy	61 GACACCAACGAAAGTGAATAAAGTCAATGAAGCCACTCCTTAAGAAAGAGTTCTCGC 120
Db	107 GACACCAACGAAAGTGAATAAAGTCAATGAAGCCACTCCTTAAGAAAGAGTTCTCGC 166
Qy	121 CGGTTTGTCACTTTTCCAATCCAGTACCTCATATTTGGAAATATGTAATAACAGGCACAG 180
Db	167 CGGTTTGTCACTTTTCCAATCCAGTACCTCATATTTGGAAATATGTAATAACAGGCACAG 226

```
Qy 181 GCTTCTCTTGGACAGCAGAGAGGTGCGATTATCAAGAGATCTCCCTCACTGGAAACAAG 240
|
Db 227 GCTTCTCTTGGACAGCAGAGAGGTGCGATTATCAAGAGATCTCCCTCACTGGAAACAAG 286
|
Qy 241 CTTAAAGCAGATGAGAGTACTTCTCTCTCACATCTTAGCTTTTGGAGCCAGTGAT 300
|
Db 287 CTTAAAGCAGATGAGAGTACTTCTCTCTCACATCTTAGCTTTTGGAGCCAGTGAT 346
|
Qy 301 GGAATTCGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTCGAGGTCTCCAGAGGCT 360
|
Db 347 GGAATTCGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTCGAGGTCTCCAGAGGCT 406
|
Qy 361 CGCTGTTTCTATGCGCTTTCAAATTTCTCATCGAGAATGTTCTCATCGAGAGATGACAGTTTG 420
|
Db 407 CGCTGTTTCTATGCGCTTTCAAATTTCTCATCGAGAATGTTCTCATCGAGAGATGACAGTTTG 466
|
Qy 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 480
|
Db 467 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 526
|
Qy 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTGGATGATGACAGATAGAAA 540
|
Db 527 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTGGATGATGACAGATAGAAA 586
|
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGATTTCTTCTCAGGA 600
|
Db 587 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGATTTCTTCTCAGGA 646
|
Qy 601 TCTTTTCTGCTATATTTCTGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
|
Db 647 TCTTTTCTGCTATATTTCTGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 706
|
Qy 661 AATGAATCTATCAGCAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 720
|
Db 707 AATGAATCTATCAGCAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 766
|
Qy 721 TACTTAGTAATAGCCTTCAGAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAA 780
|
Db 767 TACTTAGTAATAGCCTTCAGAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAA 826
|
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCCTTGCAGTGGCCCTCATTTGGAATGAATTCATT 840
|
Db 827 ATTGAGCAGGAGTTTAAACAGAGCCCTTGCAGTGGCCCTCATTTGGAATGAATTCATT 886
|
Qy 841 TTGATGAAACAGTACATGATGTTGTAGTGACAGATTAATTTGTGGAATTTGATTTCA 900
|
Db 887 TTGATGAAACAGTACATGATGTTGTAGTGACAGATTAATTTGTGGAATTTGATTTCA 946
|
Qy 901 AAGGTTTTTTCAGGAGAAATCCTTTTGTATTTATGGAACATTTCTTTAGAGGAAA 960
|
Db 947 AAGGTTTTTTCAGGAGAAATCCTTTTGTATTTATGGAACATTTCTTTAGAGGAAA 1006
|
Qy 961 ACAATTTCTTTGAGAAACAGATTTTCAGAGATATCAGGCTTTTGTGAGTTATGGCAGAAACC 1020
|
Db 1007 ACAATTTCTTTGAGAAACAGATTTTCAGAGATATCAGGCTTTTGTGAGTTATGGCAGAAACC 1066
|
Qy 1021 ACAGATAACCTCTTCACTTTGGATGCAATTTT 1053
|
Db 1067 ACAGATAACCTCTTCACTTTGGATGCAATTTT 1099
|
RESULT 3
BC042948
LOCUS Homo sapiens cDNA clone IMAGE:4798175, containing frame-shift errorB.
ACCESSION BC042948
VERSION BC042948.2 GI:34194000
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

REFERENCE
AUTHORS

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (Bases 1 to 4650)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, P.S., Wagner, K.H., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.P., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E.,
Scherer, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (Bases 1 to 4650)
Director MGC Project.
Direct Submission
Submitted (09-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 25, 2003 this sequence version replaced gi:27695575.
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabh, Parvaneh Saedi, JR Santos, Angeli Schnerch, Ursula Skalska,
Duane Smal, Jeff Stott, Miranda Tsai, George Yang, Jacques
Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAP Plate: 70 Row: h Column: 19
This clone has the following problem: frame shifted.

FEATURES
source

1. 4650
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4798175"
/tissue_type="Brain, hypothalamus"
/clone_lib="NIH MGC_96"
/lab_host="DH10B"
/note="Vector: pBluescript"

ORIGIN

Query Match 98.9%; Score 1041; DB 3; Length 4650;
Best Local Similarity 99.9%; Pred. NO. 5.8e-266;
Matches 1052; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy

1 ATGGCGGACCGGAAAGCGGAGCGCGGCTGATCAGATCAGATCAGATCCTTCA 60
|||||

Db 72 ATGGGGGACCCGGGAAAGCGCCGGGCTGGATCAGGATGAGAGATCACTTCA 131
Qy 61 GACACCAACGAAAGTAAAGTCAAAATGAAGAGCCACTCCTTAAGAAAGAGTTCTCGC 120
Db 132 GACACCAACGAAAGTAAAGTCAAAATGAAGAG-CACTCCTTAAGAAAGAGTTCTCGC 190
Qy 121 CGGTTTGTCACTTTTCCAATCAGTACCCCTGATATTTGGAAAAATGTATAAACAGGACAG 180
Db 191 CGGTTTGTCACTTTTCCAATCAGTACCCCTGATATTTGGAAAAATGTATAAACAGGACAG 250
Qy 181 GCTTCTCTTGGACAGCAGAGAGGTGCACTTATCAAGAGATCTCCCTCACTGGAACAG 240
Db 251 GCTTCTCTTGGACAGCAGAGAGGTGCACTTATCAAGAGATCTCCCTCACTGGAACAG 310
Qy 241 CTTAAAGCAGATGAGAGTACTTCACTCTCACATCTTAGCCTTTTGGCAGGACAGTAT 300
Db 311 CTTAAAGCAGATGAGAGTACTTCACTCTCACATCTTAGCCTTTTGGCAGGACAGTAT 370
Qy 301 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTTATGTCAGGAGGTGCGAGTTCCAGAGGCT 360
Db 371 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTTATGTCAGGAGGTGCGAGTTCCAGAGGCT 430
Qy 361 CGCTGTTTCTATGCTTTCAAAATTTCTCATCAGAAATGTTCACTCAGAGATGACAGTTTG 420
Db 431 CGCTGTTTCTATGCTTTCAAAATTTCTCATCAGAAATGTTCACTCAGAGATGACAGTTTG 490
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTTAATGCAATTGAA 480
Db 491 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTTAATGCAATTGAA 550
Qy 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATGGGCGCTTGGATGGATGAGATAGAAAA 540
Db 551 ACCATGCCCTATGTTAAGAAAAAAGCAGATGGGCGCTTGGATGGATGAGATAGAAAA 610
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTCTGCTGTAGAGAGGTTTCTTCTCAGGA 600
Db 611 TCTACTTTTGGGAAAGAGTGGTGGCTTCTGCTGTAGAGAGGTTTCTTCTCAGGA 670
Qy 601 TCTTTTGTGCTATATTTCTGCTAAGAGAGAGGCTTATGCCAGGACTCATTCTTTTCC 660
Db 671 TCTTTTGTGCTATATTTCTGCTAAGAGAGAGGCTTATGCCAGGACTCATTCTTTTCC 730
Qy 720 AATGAATCTATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 720
Db 731 AATGAATCTATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 790
Qy 721 TACTTAGTAAATAGCCCTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db 791 TACTTAGTAAATAGCCCTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 850
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCCTGCGAGTTCAGGCTTGGATGCAATTGCAAT 840
Db 851 ATTGAGCAGGAGTTTAAACAGAGCCCTGCGAGTTCAGGCTTGGATGCAATTGCAAT 910
Qy 841 TTGATGAAACAGTACATTTAGTGTAGCTGACAGATTTACTTGTGGAACCTTGGATTTCTCA 900
Db 911 TTGATGAAACAGTACATTTAGTGTAGCTGACAGATTTACTTGTGGAACCTTGGATTTCTCA 970
Qy 901 AAGGTTTTTCAGGAGAAAAATCCCTTTTGAATTTTATGGAACAATTTCTTTAGAGGAAAA 960
Db 971 AAGGTTTTTCAGGAGAAAAATCCCTTTTGAATTTTATGGAACAATTTCTTTAGAGGAAAA 1030
Qy 961 ACAATTTCTTTGAGAAACGAGTTTTCAGAGATCAGCGTTTGGAGTTATGGCAGAAACC 1020
Db 1031 ACAATTTCTTTGAGAAACGAGTTTTCAGAGATCAGCGTTTGGAGTTATGGCAGAAACC 1090
Qy 1021 ACAGATAAGTCTTCACTTGGATGCAATTTT 1053
Db 1091 ACAGATAAGTCTTCACTTGGATGCAATTTT 1123

LOCUS CR627376 3501 bp mRNA linear HTC 22-SEP-2004
DEFINITION Homo sapiens mRNA; cDNA DKFZp686M05248 (from clone DKFZp686M05248).
ACCESSION CR627376
VERSION CR627376.1 GI:50949847
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3501)
AUTHORS Ottenwaelder,B.; Obermaier,B., Deutschenbaue,S., Schaipp,A.,
Mewes,H.W., Weill,B., Amlid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
The German cDNA Consortium
Direct Submission
Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Medigenomix (Martinsried/Germany) within the cDNA
sequencing consortium of the German Genome Project.
This clone (DKFZp686M05248) is available at the RZPD Deutsches
Resourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp686M05248
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.
FEATURES
Location/Qualifiers
1..3501
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686M05248"
/tissue_type="salivary gland"
/clone_lib="686 (synonym: hicc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
/note="ribonucleotide reductase M2 B (TP53 inducible)"
gene
1..3501
/gene="DKFZp686M05248"
332..1225
/gene="DKFZp686M05248"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH10473.1"
/db_xref="GI:50949848"
/translation="MYKQAQASFWTAEEVDLSKDLPHWNKLKADKDYFISHTLAFPA
SDGIVNENLVERFSQEVQPEARCFYQIILIENVHSEMSILLIDTYIRDPKRRFLF
NAIETMPVKKKADWLRWIAADKSTGFRVVAFAAVGVFSGSFAAIFWLKRGFLM
PLGTFSELISRLDGLHCDPACLMFOVLVKNKPPRVEREIIIVDAVKIQEFTLEALPV
GLIGNCILMKQYIEFVADRLLELVLFSGKVFQAEPPDFMENISLEGRINFFPEKRVSE
YQRFVAVAEITDNDVFTLDADF"
ORIGIN
Query Match 96.0%; Score 1010.6; DB 3; Length 3501;
Best Local Similarity 98.6%; Pred. No. 6.8e-258;
Matches 1019; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Qy 21 GGAAGCGCGCGGCTGGATCAGGATGAGAGATCATCTTCAGACACCAACGAAAGTGAAT 80
Db 190 GGACGACGCTTGGAGCGCTCAGCGCGGAGATCATCTTCAGACACCAACGAAAGTGAAT 249
Qy 81 AAGTCAAAATGAAGAGCCACTCCTTAAGAAAGAGTTCTCGCGGTTTGTCAATCTTTCCAAT 140
Db 250 AAGTCAAAATGAAGAGCCACTCCTTAAGAAAGAGTTCTCGCGGTTTGTCAATCTTTCCAAT 309
Qy 141 CCAGTACCTGATATTTGGAATAATGTATAACAGCAGCAGGCTTCCTTCGACAGCAGA 200
Db 310 CCAGTACCTGATATTTGGAATAATGTATAACAGCAGCAGGCTTCCTTCGACAGCAGA 369
Qy 201 AGAGTCCGACTTATCAAGAGGATCTCCCTCACTGGAAACAAAGCTTAAAGCAGATGAGAAGTA 260
Db 370 AGAGTCCGACTTATCAAGAGGATCTCCCTCACTGGAAACAAAGCTTAAAGCAGATGAGAAGTA 429

Qy	661	AATGA	ACTCAT	CAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA	720
Db	661	AATGA	ACTCAT	CAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA	720
Qy	721	TACTT	AGTAAAT	AAGCCCTTCAGAGAAGGGTCAGGGAGATCATTTGTTGATGCTGTCAAA	780
Db	721	TACTT	AGTAAAT	AAGCCCTTCAGAGAAGGGTCAGGGAGATCATTTGTTGATGCTGTCAAA	780
Qy	781	ATTGAG	CAGGAGT	TTTTTAAACAGAGCCCTTGCCAGTTGGCCTCATTTGGAATGAATTGCATT	840
Db	781	ATTGAG	CAGNNNNNNNNNN	NNNNNAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTGCATT	840
Qy	841	TTGAT	TGAAA	CAGTACATGTGATTTGTGTAGCTGACAGATTACTGTGTGGAACTTTGCATCTCA	900
Db	841	TTGAT	TGAAA	CAGTACATGTGATTTGTGTAGCTGACAGATTACTGTGTGGAACTTTGCATCTCA	900
Qy	901	AAGGT	TTTTT	CAGGCAGAGAAATCCCTTTTGATTTTATGAAAACATTTCTTTAGAGGAAA	960
Db	901	AAGGT	TTTTT	CAGGCAGAAAATCCCTTTTGATTTTATGAAAACATTTCTTTAGAGGAAA	960
Qy	961	ACAAA	TTCTTT	TGAGAAAACGAGTTTCAGAGTATCAGCGTTTTCGAGTTATGCGCAGAAACC	1020
Db	961	ACAAA	TTCTTT	TGAGAAAACGAGTTTCAGAGTATCAGCGTTTTCGAGTTATGCGCAGAAACC	1020
Qy	1021	ACGAT	TAACTCTT	TTCACCTTGGATGCGAGATTTT	1053
Db	1021	ACGAT	TAACTCTT	TTCACCTTGGATGCGAGATTTT	1053

RESULT	6
AL547501	
LOCUS	
DEFINITION	1083 bp mRNA linear EST 25-MAR-2004 AL547501 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CSODI011vfl4 5-PRIME, mRNA sequence.
ACCSSION	AL547501
VERSION	AL547501.3 GI:45747948
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 1083)
REFERENCE	Li,W.B., Gruber,C., Jesse,J. and Polayes,D. Full-length cDNA libraries and normalization Unpublished (2001) On Feb 15, 2001 this sequence version replaced gi:31269332.
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

Query Match

RESULT 7

[illegible]

```

RESULT 8
BQ441857
LOCUS
DEFINITION
BQ441857 879 bp mRNA linear EST 24-MAY-2000
AGENCOURT 7627448 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6014179
5', mRNA sequence.
ACCESSION
BQ441857
VERSION
BQ441857.1 GI:21180933
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 879)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13208 row: d column: 20
High quality sequence stop: 662.
Location/Qualifiers
1. .879
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6014179"
FEATURES
source

```


/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_92"

/note="Organ: testis; Vector: pCW-SPORE6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

ORIGIN

```
Query Match      75.3%; Score 793.2; DB 5; Length 879;
Best Local Similarity 98.7%; Pred. No. 4.2e-200;
Matches 821; Conservative 0; Mismatches 8; Indels 3; Gaps 2;

Qy 1 ATGGCGCACCGGAAAGCGCGGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 22 ATGGCGCACCGGAAAGCGCGGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 81

Qy 61 GACACCAACGAAAGTGAATGAATCAATGAAGAGCCACTCTTAAAGAGAGTTCTGC 120
Db 82 GACACCAACGAAAGTGAATGAATGAATGAAGAGCCACTCTTAAAGAGAGTTCTGC 141

Qy 121 CGTTTCTGTCATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATTAACAGGCACAG 180
Db 142 CGTTTCTGTCATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATTAACAGGCACAG 201

Qy 181 GCTTCTCTTGGACAGCAGAGAGGTGACATTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db 202 GCTTCTCTTGGACAGCAGAGAGGTGACATTTATCAAGAGATCTCCCTCACTGGAACAAG 261

Qy 241 CTTAAAGCAGATGAGAGTACTTCACTCTCACATCTTAGCTTTTTCAGCAGCAGTAT 300
Db 262 CTTAAAGCAGATGAGAGTACTTCACTCTCACATCTTAGCTTTTTCAGCAGCAGTAT 321

Qy 301 GGAATTCGTAATGAAATTTGGTGAGCGCTTTAGTCAGAGAGGTTCAGAGGCT 360
Db 322 GGAATTCGTAATGAAATTTGGTGAGCGCTTTAGTCAGAGAGGTTCAGAGGCT 381

Qy 361 CGCTGTTTCTATGCTTTTCAAAATCTCATCAGAGATGTTCACTCAGAGATGTACAGTTTG 420
Db 382 CGCTGTTTCTATGCTTTTCAAAATCTCATCAGAGATGTTCACTCAGAGATGTACAGTTTG 441

Qy 421 CTGATAGACACTTACATCAGAGATGCCAAGAAAAGGGAATTTTATTTAATGCAATTTGAA 480
Db 442 CTGATAGACACTTACATCAGAGATGCCAAGAAAAGGGAATTTTATTTAATGCAATTTGAA 501

Qy 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTTCAGATGATGATGAGATAGAAA 540
Db 502 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTTCAGATGATGATGAGATAGAAA 561

Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTCTGCTGTAGAGAGATTTTCTTCTCAGGA 600
Db 562 TCTACTTTTGGGAAAGAGTGGTGGCTTTTCTGCTGTAGAGAGATTTTCTTCTCAGGA 621

Qy 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACATTTTTC 660
Db 622 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACATTTTTC 681

Qy 661 AATGAACCTCATCAGCAGATGAAGGACTTCACTGTGACTTTTGTGCTGATGTTCCAA 720
Db 682 AATGAACCTCATCAGCAGATGAAGGACTTCACTGTGACTTTTGTGCTGATGTTCCAA 741

Qy 721 TACTTAGTAAATAGCCCTCAGAGAAAGGGG--TCAGGGAGATCATTTGTTGATGCTGTCA 778
Db 742 TACTTAGTAAATAGCCCTTCAAAAGAAAGGGGTCAGGGGAGATCATTTGTTGATGCTGTCA 801

Qy 779 AAATT--GAGCAGGAGTATTTTAAACAGAAAGCCCTTGCAGTTGGCCCTCATTTGAA 829
Db 802 AAATTGGACGAGGATTTTAAACAGAAAGCCCTTGCAGTTGGCCCTCAATTGGA 853
```

RESULT 9

AY398975

LOCUS

1043 bp DNA linear GSS 12-DEC-2003

DEFINITION Mus musculus HCM0069 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.

ACCESSION AY398975

VERSION AV398975.1 GI:39754964

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 1043)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 1043)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.

FEATURES
source
1..1043
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1043
/locus_tag="HCM0069"

gene

ORIGIN

```
Query Match      73.8%; Score 777; DB 9; Length 1043;
Best Local Similarity 81.2%; Pred. No. 9.2e-196;
Matches 854; Conservative 0; Mismatches 185; Indels 13; Gaps 1;

Qy 1 ATGGCGCACCGGAAAGCGCGGAGCGCGGGCTGGATCAGAGATGAGAGATCATCTTCA 60
Db 1 ATGGCGCACCGGAAAGCGCGGAGCGCGGGCTGGATCAGAGATGAGAGATCATCTTCA 60

Qy 61 GACACCAACGAAAGTGAATGAATGAAGAGCCACTCTTAAAGAGAGTTCTCCTC 120
Db 61 GACACCAACGAAAGTGAATGAATGAAGAGCCACTCTTAAAGAGAGTTCTCCTC 120

Qy 121 CGGTTTGTCTATCTTTTCCAAATCCAGTACCTGATATTTTGGAAATGTATTAACAGGCACAG 180
Db 121 CGGTTTGTCTATCTTTTCCAAATCCAGTACCTGATATTTTGGAAATGTATTAACAGGCACAG 180

Qy 181 GCTTCTCTTGGACAGCAGAGAGGTGACATTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db 181 GCTTCTCTTGGACAGCAGAGAGGTGACATTTTATCAAGAGATCTCCCTCACTGGAACAAG 240

Qy 241 CTTAAAGCAGATGAGAGTACTTCACTCTCACATCTTAGCTTTTTCAGCAGCAGTAT 300
Db 241 CTTAAATCCGACGAGAGATATTTTATCTTCCACATCTTAGCTTTTTCAGCAGCAGTAT 300

Qy 301 GGAATTCGTAATGAAATTTGGTGAGCGCTTTAGTCAGAGAGGTTCAGAGGCT 360
Db 301 GGAATTCGTAATGAAATTTGGTGAGCGCTTTAGTCAGAGAGGTTCAGAGGCT 360

Qy 361 CGCTGTTTCTATGCTTTTCAAAATCTCATCAGAGATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGCTTTTCAAAATCTCATCAGAGATGTTCACTCAGAGATGTACAGTTTG 420

Qy 421 CTGATAGACACTTACATCAGAGATGCCAAGAAAAGGGAATTTTATTTAATGCAATTTGAA 480
Db 421 CTAATAGACACTTACATCAGAGATGCCAAGAAAAGGGAATTTTATTTAATGCAATTTGAA 480
```


	Accession	(Long) Homo sapiens cDNA clone IMAGE:30427288 5', mRNA sequence.
	Version	CD656661
	Keywords	EST, CD656661.1 GI:31897180
	Source	Homo sapiens (human)
	Reference	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 804) NTH-MGC http://mgc.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Daniela S. Gerhard, Ph.D. Office of Cancer Genomics National Cancer Institute / NIH Bldg. 31 Rm10A07 Bethesda, MD 20892 Email: cgabs@emil.nih.gov Tissue procurement: Irene Ginis and Mahendra Rao, NIA cDNA Library Preparation: Yulan Piao and Minoru KO cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC c lone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: NDAM514 row: h column: 17 High quality sequence stop: 680.
	Title	National Institutes of Health, Mammalian Gene Collection (MGC)
	Journal	Unpublished (1999)
	Comment	Contact: Daniela S. Gerhard, Ph.D. Office of Cancer Genomics National Cancer Institute / NIH Bldg. 31 Rm10A07 Bethesda, MD 20892 Email: cgabs@emil.nih.gov Tissue procurement: Irene Ginis and Mahendra Rao, NIA cDNA Library Preparation: Yulan Piao and Minoru KO cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC c lone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: NDAM514 row: h column: 17 High quality sequence stop: 680.
	Features	1..804 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:30427288" /tissue_type="Embryonic Stem cells" /cell_line="WA01" /lab_host="DH10B (T1 phage-resistant)" /clone_lib="NIA Human Hi Embryonic Stem Cell cDNA Library (Long)" /notes="Vector: pCMW-Sport6; Site 1: NotI; Site 2: SalI; This is a long-transcript enriched cDNA library (Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01 cell line. Undifferentiated human ES cell line WA01/HI was obtained from WiCell Research Institute, Inc., Madison, WI, cultured according to their instructions, on MEK feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSEA-4, OCT3, OCT4, REVI, UTP, TERT, SOX2, CX43 and CX45. They are negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1, TUBB3, NES, GFAP, and BOMES. When confluent (18-10 days after plating), the ES cells from 4 x 6 cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/GIBCO) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with TRIzol Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID:11544199]] Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTCCTAGATCGCAGCGCCGCTTTTCTTTT-3'] from 3.4g of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform extraction, and separated from free linkers by Centricon-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S for 25 cycles. The product were purified by phenol/chloroform extraction and Centricon-100 column. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMW-SPORT6 plasmid vector. The average insert size is about 3.6kb."
	Source	1..804 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:30427288" /tissue_type="Embryonic Stem cells" /cell_line="WA01" /lab_host="DH10B (T1 phage-resistant)" /clone_lib="NIA Human Hi Embryonic Stem Cell cDNA Library (Long)" /notes="Vector: pCMW-Sport6; Site 1: NotI; Site 2: SalI; This is a long-transcript enriched cDNA library (Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01 cell line. Undifferentiated human ES cell line WA01/HI was obtained from WiCell Research Institute, Inc., Madison, WI, cultured according to their instructions, on MEK feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSEA-4, OCT3, OCT4, REVI, UTP, TERT, SOX2, CX43 and CX45. They are negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1, TUBB3, NES, GFAP, and BOMES. When confluent (18-10 days after plating), the ES cells from 4 x 6 cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/GIBCO) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with TRIzol Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID:11544199]] Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTCCTAGATCGCAGCGCCGCTTTTCTTTT-3'] from 3.4g of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform extraction, and separated from free linkers by Centricon-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S for 25 cycles. The product were purified by phenol/chloroform extraction and Centricon-100 column. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMW-SPORT6 plasmid vector. The average insert size is about 3.6kb."

RESULT	13
CD656661	
LOCUS	CD656661
DEFINITION	AGENCOURT_1455875 NIA Human H1 Embryonic Stem Cell cDNA library
	804 bp mRNA linear EST 18-JUN-2003

ORIGIN

Query Match 64.8%; Score 682.8; DB 6; Length 804;
Best Local Similarity 97.3%; Pred. No. 1.1e-170;
Matches 716; Conservative 0; Mismatches 17; Indels 3; Gaps 2;

```
Qy 212 TATCAAGGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTACTTCTCTC 271
Db 1 TATCAAGGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTACTTCTCTC 60
Qy 272 ACATCTTAGC-CTTTTTCAGCCAGTATGGAATGTAATGAAAATTTGGTGGAGGC 330
Db 61 ACATCTTAGC-CTTTTTCAGCCAGTATGGAATGTAATGAAAATTTGGTGGAGGC 120
Qy 331 TTTAGTCAGGAGTGCAGGTTCCAGAGCTCGCTGTTTCTATGCTTTCAAATTTCTCATC 390
Db 121 TTTAGTCAGGAGTGCAGGTTCCAGAGCTCGCTGTTTCTATGCTTTCAAATTTCTCATC 180
Qy 391 GAGATGTTCTCACTCAGAGATGTACAGTTTGTCTGTATGACACTTATCATCAGAGATCCCAAG 450
Db 181 GAGATGTTCTCACTCAGAGATGTACAGTTTGTCTGTATGACACTTATCATCAGAGATCCCAAG 240
Qy 451 AAAAGGGAATTTTATTTAATGCAATTAACCAATGCAATGCAATGCAATGCAATGCAATGCAAT 510
Db 241 AAAAGGGAATTTTATTTAATGCAATTAACCAATGCAATGCAATGCAATGCAATGCAATGCAAT 300
Qy 511 TGGCCCTTGGCATGATGATAGAGAAATCTACTTTTGGGAAAGAGTGGTGGCCCTT 570
Db 301 TGGCCCTTGGCATGATGATAGAGAAATCTACTTTTGGGAAAGAGTGGTGGCCCTT 360
Qy 571 GCTGCTGTAGAGAGATTTTCTTCTCAGGATCTTTTGTCTGTATATTTCTGCTTAAAGAG 630
Db 361 GCTGCTGTAGAGAGATTTTCTTCTCAGGATCTTTTGTCTGTATATTTCTGCTTAAAGAG 420
Qy 631 AGAGTCTTATGCGAGACTCACTTTTCCATGACTCATCAGCAGAGATGAAGACTT 690
Db 421 AGAGTCTTATGCGAGACTCACTTTTCCAACTCATCAGCAGAGATGAAGACTT 480
Qy 691 CACTGTGACTTGTCTGCTGATGCTTCCAACTTACTTAAATAGCCCTTCAAGAGAGG 750
Db 481 CACTGTGACTTGTCTGCTGATGCTTCCAACTTACTTAAATAGCCCTTCAAGAGAGG 540
Qy 751 GTACAGGAGATCAATGTTGATGCTGTCAAAATTCAGCAGAGATTTTAAACAGAGCCCTG 810
Db 541 GTACAGGAGATCAATGTTGATGCTGTCAAAATTCAGCAGAGATTTTAAACAGAGCCCTG 600
Qy 811 CCAGTGGCCCTCATGGAATGCAATTTGATGAAACAGTACATGATGATTTGTAGCT 870
Db 601 CCAGTGGCCCTCATGGAATGCAATTTGATGAAACAGTACATGATGATTTGTAGCT 660
Qy 871 GACAGATCTGTTGGAACTTGGATTTCTCAAGGTTTTCAGGAGAGAAATCCTTTTGTAT 930
Db 661 GACAGATCTGTTGGAACTTGGAT--CTCAAGGTTTTCAGGAGAGAAATCCTTTTGTAT 718
Qy 931 TTTATGGAACAATTT 946
Db 719 TATGGAACAATCTTT 734
```

```
RESULT 14
BM801298
LOCUS BM801298
DEFINITION AGENCOURT_6422860 NIH_MGC_92 Homo sapiens cdna clone IMAGE:5586693
5', mRNA sequence.
ACCESSION BM801298
VERSION BM801298.1 GI:19118121
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1061)
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
CONTACT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
```

cdna Library Preparation: Life Technologies, Inc.
cdna Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
<http://image.llnl.gov>
Plate: LLAM12354 row: p column: 22
High quality sequence stop: 631.

FEATURES

source
1..1061
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5586693"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 62.2%; Score 655.4; DB 4; Length 1061;
Best Local Similarity 97.6%; Pred. No. 2.4e-163;
Matches 718; Conservative 0; Mismatches 12; Indels 6; Gaps 5;
Qy 1 ATGGGCGACCCGGAAGCCCGGAGCGCCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 85 ATGGGCGACCCGGAAGCCCGGAGCGCCGGCTGGATCAGGATGAGAGATCATCTTCA 144
Qy 61 GACACCAACGAAGTGAATAAAGTCAANTCAAGAGCCACTCTTAAGAAAGAGTTCTGC 120
Db 145 GACACCAACGAAGTGAATAAAGTCAANTCAAGAGCCACTCTTAAGAAAGAGTTCTGC 204
Qy 121 CGGTTTGTCTATCTTTTCCAATCCAGTACCTCTGATATTTGGAAAATGTATAACAGGCACAG 180
Db 205 CGGTTTGTCTATCTTTTCCAATCCAGTACCTCTGATATTTGGAAAATGTATAACAGGCACAG 264
Qy 181 GCTTCCTTCTGGACAGCAGAGAGTGCATCTTATCAAGAGATCTCCCTCATCTGGAACAG 240
Db 265 GCTTCCTTCTGGACAGCAGAGAGTGCATCTTATCAAGAGATCTCCCTCATCTGGAACAG 324
Qy 241 CTTAAAGCAGATGAGAGTACTTCTATCTCTCACATCTTAGCTTTTTCAGCCAGTGTAT 300
Db 325 CTTAAAGCAGATGAGAGTACTTCTATCTCTCACATCTTAGCTTTTTCAGCCAGTGTAT 384
Qy 301 GGAATTGTAAATGAAAATTTGCTGGAGCGCTTTAGTCAGAGAGGTGCAGAGTTCAGAGGCT 360
Db 385 GGAATTGTAAATGAAAATTTGCTGGAGCGCTTTAGTCAGAGAGGTGCAGAGTTCAGAGGCT 444
Qy 361 CGCTGTTTCTATGCTTTCAAATTTCTATCAGAAATGTTCTCATCAGATGTACAGTTTG 420
Db 445 CGCTGTTTCTATGCTTTCAAATTTCTATCAGAAATGTTCTCATCAGAGATGTACAGTTTG 504
Qy 421 CTGTATGACACTTATCATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATGGA 480
Db 505 CTGTATGACACTTATCATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATGGA 564
Qy 481 ACCATGCTCTATGTTTAAAGAAAAGCAGATTTGGGCTTTGCGATGATGATAGCAGATAGAAA 540
Db 565 ACCATGCTCTATGTTTAAAGAAAAGCAGATTTGGGCTTTGCGATGATGATAGCAGATAGAAA 624
Qy 541 TCTACTTTTGGGAAAAGAGTGGTGGCCCTTCTGCTGTAGAAAGAGTTTCTTCTCAGGA 600
Db 625 TCTACTTTTGGGAAAAGAGTGGTGGCCCTTCTGCTGTAGAAAGAGTTTCTTCTCAGGA 684
Qy 601 TCTTTTGTCTGTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCATCTTTTTC 660
Db 685 TCTTTTGTCTGTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCATCTTTTTC 743
Qy 661 AATGAACCTCAT-CAGCAGAGAT-GAAGGACTTCACTGTGA--CTTTGCTTGGCTCAT-GT 715

```

Db      744 CATGACTCATCCACGAGATGAGGAGCTTCCCTGTGAACCTTGCCTTGCCTGATGCT 803
Qy      716 TCCAATATCTTAGTAA 731
Db      804 TCCAATACTNAGAA 819

RESULT 15
BUI70979
LOCUS   BUI70979
DEFINITION AGENCOURT_7940338 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6143754
5', mRNA sequence.
BUI70979
VERSION BUI70979.1 GI:22684963
KEYWORDS EST.
SOURCE  Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 896)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-x@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13467 row: k column: 19
High quality sequence stop: 619.

FEATURES             source
    1..896
        Location/Qualifiers
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:6143754"
            /tissue_type="retinoblastoma"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_67"
            /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
            Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
            Average insert size 1.75 kb. Library constructed by Life
            Technologies."

ORIGIN

Query Match      62.2%; Score 654.8; DB 5; Length 896;
Best Local Similarity 98.7%; Pred. No. 3.4e-163;
Matches 671; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

Qy      1 ATGGGGACCCGGAAGCGCGCGGCTGATCAGATGAGATCATCTTCA 60
Db      55 ATGGGGACCCGGAAGCGCGGAGCGCGGCTGATCAGATGAGATCATCTTCA 114
Qy      61 GACACCAAGAAAGTGAATTAAGTCAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 120
Db      115 GACACCAAGAAAGTGAATTAAGTCAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 174
Qy      121 CGGTTTGTCTATCTTCCATCCAGTACCTCGATATTTGGAAATGTATATAACAGGCACAG 180
Db      175 CGGTTTGTCTATCTTCCATCCAGTACCTCGATATTTGGAAATGTATATAACAGGCACAG 234
Qy      181 GCTTCTCTTGGACAGCAGAGAGGTGCGATTATCAAGGATCTCCCTCACTGGAACAAG 240
Db      235 GCTTCTCTTGGACAGCAGAGAGGTGCGATTATCAAGGATCTCCCTCACTGGAACAAG 294
Qy      241 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCAATCTTAGCCTTTTTCAGCCAGTGAT 300
Db      295 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCAATCTTAGCCTTTTTCAGCCAGTGAT 354

```

```

Qy      301 GGAATTGTAAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCAGAGGCT 360
Db      355 GGAATTGTAAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCAGAGGCT 414
Qy      361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 420
Db      415 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 474
Qy      421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db      475 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 534
Qy      481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCTTTCGGATGATAGCAGATACAAAA 540
Db      535 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCTTTCGGATGATAGCAGATACAAAA 594
Qy      541 TCTACTTTTGGGGAAGAGTGGTGGCTTTTGTCTGTGTAGAAGGAGTTTCTTCTCAGGA 600
Db      595 TCTACTTTTGGGGAAGAGTGGTGGCTTTTGTCTGTGTAGAAGGAGTTTCTTCTCAGGA 654
Qy      601 TCTTTTGTCTATATTTCTGGCTAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 658
Db      655 TCTTTTGTCTATATTTCTGGCTAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 714
Qy      659 CCAATGAACCTCATCAGCAGA 678
Db      715 CCAATGAACCTCATCAGCAGA 734

```

Search completed: September 26, 2005, 02:21:58
Job time : 2741 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 1, 2005, 12:09:09 ; Search time 0.001 Seconds
(without alignments)
1478.412 Million cell updates/sec

Title: US-10-698-228-1
Perfect score: 1821
Sequence: 1 MGDPERPEAGLDQDERSSS.....QRFVMAETDNNVFTLDADF 351

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2 seqs, 2106 residues

Total number of hits satisfying chosen parameters: 4

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DRV=soft -O=seq1.rep -DB=seq -SUFFIX=pto
-OUT=swope228.res -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1
-END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALLIGN=200
-THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -NCPU=6 -NO XLPXY
-NEG SCORES=0 -LONGIOG -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOF=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : *seq:*
1: /home/tport/1005/swope228/align/seq2.seq*
2: /home/tport/1005/swope228/align/seq12.seq*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	100.0	1053	1	US-10-698-228-2
2	1821	100.0	1053	2	US-10-698-228-2
3	41.5	2.3	1053	1	US-10-698-228-2
4	38.5	2.1	1053	2	US-10-698-228-12

ALIGNMENTS

RESULT 1
US-10-698-228-2
; Sequence 2, Application US/10698228
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28

; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2

; LENGTH: 1053

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION:

US-10-698-228-2

Alignment Scores:

Pred. No.:	Score:	Length:
0	1821.00	1053
Percent Similarity:	100.00%	Matches: 351
Best Local Similarity:	100.00%	Conservative: 0
Query Match:	100.00%	Mismatches: 0
DB:	1	Indels: 0
		Gaps: 0

US-10-698-228-1 (1-351) x US-10-698-228-2 (1-1053)

Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20

Db 1 ATGGCGCACC CGGAAGCGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60

Qy 21 AspThrAsnGluSerGluLeuLysSerAsnGluProLeuLeuArgLysSerArg 40

Db 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120

Qy 41 ArgPheValIlePhePheProIleGluTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60

Db 121 CGGTTTGTCTATCTTCCAATCAGTACCTCATATTGGAAATATGATAAAGGACAG 180

Qy 61 AlaSerPheTyrThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80

Db 181 GCTTCCTCTGGACAGCAGAGAGAGTGCATTTATCAAGAGATCTCCCTCAGTGGAAACAAG 240

Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAlaSerAsp 100

Db 241 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTTAGCCTTTTTCAGCCAGTGTAT 300

Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120

Db 301 GGAATTGTAAATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360

Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140

Db 361 CGCTGTTTCTATGGCTTTCAAAATTTCTATCAGAGATGTTTCTCAGAGATGTACAGTTTG 420

Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160

Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTATTAATCAATTGAA 480

Qy 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180

Db 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCGCTTTCGATGGATGACAGATGAAAA 540

Qy 181 SerThrPheGlyGluArgValValAlaPheAlaValGluGluValPhePheSerGly 200

Db 541 TCTACTTTTGGGAAAGAGTGTGGCTTTCTCTGTAGAGAGGAGTTTCTTCTCAGGA 600

Qy 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220

Db 601 TCTTTTGTCTATATTTCTGGCTAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660

Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240

Db 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTGTCTTGTCTGTATGTTCCAA 720

QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAlaAspAlaValLys 260
DB 721 TACTTAGTAAATAGACCTTCAGAGAAAGGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 781 ATTGAGCAGGAGTTTTTAACAGAGCCTTGCCAGTTGGCTCATTTGGAAATGATTCATT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGTGGAACCTTGGATTCTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 901 AAGGTTTTTCAGGCAGAAAATCCTTTGATTTTATGGAATAACATTTCTTTAGAAAGAAA 960
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 961 ACAAATTTCTTTGAGAAACGAGTTTCAGAGATATCAGCGTTTTCAGAGTTATGCGAGAAACC 1020
QY 341 ThrAsnValPheThrLeuAspAlaAspPhe 351
DB 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053

RESULT 2

US-10-698-228-12
; Sequence 12, Application US/10698228
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12

LENGTH: 1053

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION:

US-10-698-228-12

Alignment Scores:

Pred. No.:	0	Length:	1053
Score:	1821.00	Matches:	351
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-10-698-228-1 (1-351) x US-10-698-228-12 (1-1053)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 1 ATGGGCGACCCGGAAAGCCGAGCGGCGGGCTGGATCAGATGAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 61 GACACCAACGAAAGTCAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
DB 121 CGGTTTGTTCATCTTCCCAATCCAGTACCTGATTTATTTGGAAAATGATTAACAGGCACAG 180
QY 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80

DB 181 GCTTCTCTTGGACAGCAGAAAGGTTGACTTATCAAGAGATCTCCCTCATCTGGAACAAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
DB 241 CTTAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCCTTTTTCAGCAGCATGAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTTTCTATGGCTTTCAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACATTTACATCAGATGCCAGAAAGGAAATTTTATTTAATGCAATTTGAA 480
QY 161 ThrMetProTyrValLysLysAlaAspTyrPalalaLeuArgTyrIleAlaAspArgLys 180
DB 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCTTGGATGATAGCAGATAGAAAA 540
QY 181 SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
DB 541 TCTACTTTTGGGAAAGAGTGGGCTTTGCTGCTAGAAAGAGTTTCTTCTCAGGA 600
QY 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 601 TCTTTTGTGCTATATTTCTGGCTTAAGAGAGAGGTCTTATGCCAGGACTCCTCTTTTCC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 661 AATGAACTCATCAGCAGATGAAGACTTCACTGTGACTTTGCTTGTGCTGCTGCTTCCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB 721 TACTTAGTAAATAGGCTTCAGAAAGAAAGGCTCAGGAGATCATTTGTTGCTGTCTCAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 781 ATTGAGCAGGAGTTTTTAAACAGAGCCCTTGCAGTTGGCCTCATTTGGAATGAAATGCATT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGTGGAACCTTGGATTTCTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 901 AAGGTTTTTCAGGCAGAAAATCCTTTTGAATTTTATGGAATAACATTTCTTTTGAAGGAAA 960
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 961 ACAAATTTCTTTGAGAAACGAGTTTTCAGAGATATCAGCGTTTTCAGAGTTATGCGAGAAACC 1020
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053

RESULT 3

US-10-698-228-2/c
; Sequence 2, Application US/10698228
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391

```
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-2
Alignment Scores:
Pred. No.: 0 Length: 1053
Score: 41.50 Matches: 19
Percent Similarity: 47.17% Conservative: 6
Best Local Similarity: 35.85% Mismatches: 23
Query Match: 2.28% Indels: 6
DB: 1 Gaps: 3
US-10-698-228-1 (1-351) x US-10-698-228-2 (1-1053)
Qy 77 HisTriP---AsnLysLeuLysAlaAspGluLysTy-PheIleSerHisIleLeuAlaPhe 95
Db 298 CACTGGCTGCAAAAAGGCTTAAGATGTGAGAGATGAAGTACTTCTCATCTGCTT---TAA 242
Qy 96 PheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluVal 115
Db 241 GCTTGTTCAGTGAGGGA-----GATCCTTTGATAGTGCACC-TCTTCTGCTGTC 192
Qy 116 GlnValProGluAlaArgCysPheTyGlyPheGlnIle 128
Db 191 CAGAAGGAAGCCGTGCTGCTGTTTATACATTTTCCAAATA 153
RESULT 4
US-10-698-228-12/c
; Sequence 12, Application US/10698228
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-12
Alignment Scores:
Pred. No.: 0 Length: 1053
Score: 38.50 Matches: 30
Percent Similarity: 36.17% Conservative: 21
Best Local Similarity: 21.28% Mismatches: 63
Query Match: 2.11% Indels: 27
DB: 2 Gaps: 6
US-10-698-228-1 (1-351) x US-10-698-228-12 (1-1053)
Qy 198 PheSerGlySerPheAla-----AlaIlePheTrpLeuLysLysArgGlyLeu 213
Db 977 TTCTCAAGAAATTTGTTTTTCTTAAAGAAATGTTTCCATAAAATCAAAGGATTT 918
```

```
Qy 214 MetProGlyLeuThrPheSerAsn-----GluLeuIleSerArgaspGluGly 229
Db 917 TCTGCCTGAAAAACCTTTGAGAAATCCAAAGTTCCACAGTAATCTGTCTAGCTACAACTCA 858
Qy 230 LeuHisCysAspPheAlaCysLeuMetPheGlnTyLeuValAsnLysProSerGluGlu 249
Db 857 ATGTACTGT-----TTTCATCAAAATGCAATTCATTCCAATGAGGCCAACTGGCAAG 807
Qy 250 ArgValArgGluIleIleValAspAlaValLysIleGluGlnGluPheLeuThrGluAla 269
Db 806 GCT-----TCTGTTAAAAAATCCTGCTCAATTTTGACAGCATCA 768
Qy 270 LeuProValGlyLeuIleGlyMetAsnCysIleLeuMetLysGlnTyIleGluPheVal 289
Db 767 ACAATGATCTCCCTGACCCCTTTCTTCTGAAGGCTTATTACTAAGTATTGGAACATCAGG 708
Qy 290 -----AlaAspArgLeuLeuValGluLeuGlyPheSerLysValPhe 303
Db 707 CAAGCAAAAGTCACAGTGAAGTCCTTCTCATCTCTGCTGATGAGTTTCGAAAAAAGTACT 648
Qy 304 GlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeu-----GluGlyLys 320
Db 647 CCTGGCATAGACCTTCTTCTTAGCCGGAATATAGCAGAAAGATCCTCGAGAGAAA 588
Qy 321 Thr 321
Db 587 ACT 585
```

Search completed: November 1, 2005, 12:09:10
Job time : 1 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2005, 06:35:16 ; Search time 74 Seconds
(without alignments)
1834.499 Million cell updates/sec

Title: US-10-698-228-1
Perfect score: 1821
Sequence: 1 MGDPERPEAAGLDQDERSSS.....QRFVAWETDNNVTLADP 351

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04: *
1: geneseqp1980s: *
2: geneseqp1990s: *
3: geneseqp2000s: *
4: geneseqp2001s: *
5: geneseqp2002s: *
6: geneseqp2003as: *
7: geneseqp2003bs: *
8: geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1821	100.0	351	3	AAY84439 Amino aci
2	1821	100.0	351	4	AAB69050 Human rib
3	1817	99.8	351	4	AAB93497 Human pro
4	1454	79.8	389	4	AAB69051 Human R2
5	1454	79.8	389	6	ABU07433 Protein d
6	1454	79.8	389	7	ADJ68753 Human hea
7	1454	79.8	389	8	ADJ66564 Ribonucle
8	1454	79.8	389	8	ADK70435 Respirato
9	1454	79.8	389	8	ADN03789 Antipsori
10	1454	79.8	389	8	ADN04444 Antipsori
11	1454	79.8	389	8	ADN019226 Human PRO
12	1454	79.8	389	8	ADQ09272 Human RRM
13	1454	79.8	389	8	ABM80258 Tumour-as
14	1454	79.8	413	3	AAB43302 Human can
15	1454	79.8	453	4	AU28017 Human con
16	1454	79.8	453	4	ABG15287 Novel hum
17	1429	78.5	390	2	AAW41765 Human rib
18	1425.5	78.3	386	8	ADQ057309 Kidney de
19	1348	74.0	393	4	ABE64130 Drosophi
20	1242	68.2	430	8	ADN21469 Bacteri
21	1230	67.5	405	6	ABJ26187 Aspergill
22	1214	66.7	381	8	ADN23033 Bacteri
23	1192.5	65.5	381	6	ABJ25587 Aspergill
24	1192	65.5	413	5	ABP73700 Candida a
25	1176.5	64.6	321	8	ADS21332 Bacteri

26	1168.5	64.2	399	6	ABR53196 Protein s
27	1168.5	64.2	399	7	ADK63240 Disease t
28	1168.5	64.2	399	8	ADS43740 Bacterial
29	1161.5	63.8	391	8	ADS44135 Bacterial
30	1161.5	63.8	401	7	ADB70101 C. neofo
31	1130	62.1	329	3	AAG51528 Arabidops
32	1130	62.1	329	3	AAG05699 Arabidops
33	1130	62.1	332	3	AAG51527 Arabidops
34	1130	62.1	332	3	AAG05698 Arabidops
35	1130	62.1	332	3	AAG05697 Arabidops
36	1117	61.3	347	3	AAG14546 Arabidops
37	1112.5	61.1	330	3	AAG14548 Arabidops
38	1112.5	61.1	333	3	AAG14547 Arabidops
39	1099	60.4	314	3	AAG51529 Arabidops
40	1091	59.9	341	3	AAG41395 Arabidops
41	1091	59.9	363	3	AAG41394 Arabidops
42	1084.5	59.6	330	3	AAG41396 Arabidops
43	1072.5	58.9	274	2	AAW26418 Swinepox
44	1072.5	58.9	274	4	AAB68239 Protein e
45	977.5	53.7	322	2	AAR72736 Plasmodiu

ALIGNMENTS

RESULT 1		AY84439 standard; protein; 351 AA.	
AY84439	AC	AC	
XX	XX	XX	
DT	DT	DT	
XX	XX	XX	
DE	25-JUL-2000 (first entry)		
XX	Amino acid sequence of a human RNA-associated protein.		
KW	Human; RNA-associated protein; cell proliferation; cancer; inflammation; immune response; reproductive disorder; actinic keratosis;		
KW	atherosclerosis; arteriosclerosis; bursitis; cirrhosis; hepatitis;		
KW	mixed connective tissue disease; myelofibrosis; primary thrombocytopenia;		
KW	paroxysmal nocturnal hemoglobinuria; polycythemia vera; psoriasis;		
XX	trauma.		
XX	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	Modified-site 18	/note= "potential phosphorylation site"	
FT	Modified-site 23	/note= "potential glycosylation site"	
FT	Modified-site 29	/note= "potential phosphorylation site"	
FT	Modified-site 38	/note= "potential phosphorylation site"	
FT	Modified-site 39	/note= "potential phosphorylation site"	
FT	Modified-site 65	/note= "potential phosphorylation site"	
FT	Modified-site 87	/note= "potential phosphorylation site"	
FT	Modified-site 182	/note= "potential phosphorylation site"	
FT	Modified-site 225	/note= "potential phosphorylation site"	
FT	Modified-site 314	/note= "potential phosphorylation site"	
FT	Modified-site 329	/note= "potential glycosylation site"	
FT	Modified-site	/note= "potential phosphorylation site"	
XX	WO200015799-A2.		
PN	23-MAR-2000.		
PD	17-SEP-1999;		
PF	99WO-US021688.		

XX 17-SEP-1998; 98US-00156039.
 PR 22-SEP-1998; 98US-00158720.
 PR 04-NOV-1998; 98US-00186815.
 PR 08-APR-1999; 99US-0128660F.
 XX (INCY-) INCYTE PHARM INC.
 XX PA
 XX Tang YT, Corley NC, Guegler KJ, Gorgone GA, Patterson C;
 PI Hillman JL, Baughn MR, Lal P, Azimzai Y, Yue H, Yang J;
 XX WPI; 2000-271437/23.
 DR N-PSDB; AAA12411.
 XX New polypeptides and polynucleotides, useful for preventing and treating
 PT a disorder associated with increased or decreased expression of RNA
 PT associated proteins.
 XX Claim 1; Page 103-104; 131pp; English.
 XX The present sequence represents a human RNA-associated protein. The
 CC expression of RNA-associated proteins is closely associated with
 CC reproductive tissues, nervous tissues, cell proliferation including
 CC cancer, inflammation and immune responses, and so they may be used for
 CC diagnosis, treatment or prevention of cell proliferative,
 CC immune/inflammatory disorders, and reproductive disorders. Diseases and
 CC disorders which may be treated include actinic keratosis,
 CC atherosclerosis, arteriosclerosis, bursitis, cirrhosis, hepatitis, mixed
 CC connective tissue disease, myelofibrosis, paroxysmal nocturnal
 CC hemoglobinuria, polycythemia vera, psoriasis, primary thrombocytopenia
 CC and cancers, and trauma
 XX SQ Sequence 351 AA;
 Query Match 100.0%; Score 1821; DB 3; Length 351;
 Best Local Similarity 100.0%; Pred. No. 5.2e-179;
 Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGDPERPAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWKMYKQAO 60
 Db 1 MGDPERPAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWKMYKQAO 60
 QY 61 ASFWTAEVDLSKDLPHWNKLKADSKYFISHILAFPAASDGI VNNLVERFSQEVQVPEA 120
 Db 61 ASFWTAEVDLSKDLPHWNKLKADSKYFISHILAFPAASDGI VNNLVERFSQEVQVPEA 120
 QY 121 RCFYGFQIILIENVHSEMYSLIIDTYIRDPKKREFLFNAIETMPYVKKKADWALRWIADRK 180
 Db 121 RCFYGFQIILIENVHSEMYSLIIDTYIRDPKKREFLFNAIETMPYVKKKADWALRWIADRK 180
 QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDFACLMPQ 240
 Db 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDFACLMPQ 240
 QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVGLGFS 300
 Db 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVGLGFS 300
 QY 301 KVFQAEENPFDPMENISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF 351
 Db 301 KVFQAEENPFDPMENISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF 351
 RESULT 2
 AAB69050
 ID AAB69050 standard; protein; 351 AA.
 XX AAB69050;
 AC AAB69050;
 XX 18-APR-2001 (first entry)
 DT 18-APR-2001 (first entry)
 XX Human ribonucleotide reductase TP53R2H protein sequence SEQ ID NO:1.
 DE Human ribonucleotide reductase; cancer; DNA repair; p53.
 XX Homo sapiens.
 XX WO200100799-A1.
 XX 04-JAN-2001.
 XX 27-JUN-2000; 2000WO-JP004189.
 XX 28-JUN-1999; 99JP-00181131.
 PR 06-JUL-1999; 99JP-00192391.
 PR 21-JAN-2000; 2000JP-00017770.
 XX (TAKE) TAKEDA CHEM IND LTD.
 PA (NAKA/) NAKAMURA Y.
 XX Nakamura Y, Arakawa H, Tanaka H;
 PI WPI; 2001-112446/12.
 DR N-PSDB; AAF32438.
 XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
 PT diagnosis, treatment and prevention of cancer.
 XX Claim 1; Fig 1-3; 102pp; Japanese.
 XX The present sequence represents a human ribonucleotide reductase
 CC designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of
 CC the DNA repair mechanism and its activity is induced by p53. It can be
 CC used for the treatment, prevention and diagnosis of a wide range of
 CC cancers
 XX SQ Sequence 351 AA;
 Query Match 100.0%; Score 1821; DB 4; Length 351;
 Best Local Similarity 100.0%; Pred. No. 5.2e-179;
 Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGDPERPAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWKMYKQAO 60
 Db 1 MGDPERPAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWKMYKQAO 60
 QY 61 ASFWTAEVDLSKDLPHWNKLKADSKYFISHILAFPAASDGI VNNLVERFSQEVQVPEA 120
 Db 61 ASFWTAEVDLSKDLPHWNKLKADSKYFISHILAFPAASDGI VNNLVERFSQEVQVPEA 120
 QY 121 RCFYGFQIILIENVHSEMYSLIIDTYIRDPKKREFLFNAIETMPYVKKKADWALRWIADRK 180
 Db 121 RCFYGFQIILIENVHSEMYSLIIDTYIRDPKKREFLFNAIETMPYVKKKADWALRWIADRK 180
 QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDFACLMPQ 240
 Db 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDFACLMPQ 240
 QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVGLGFS 300
 Db 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVGLGFS 300
 QY 301 KVFQAEENPFDPMENISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF 351
 Db 301 KVFQAEENPFDPMENISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF 351
 RESULT 3
 AAB93497
 ID AAB93497 standard; protein; 351 AA.
 XX AAB93497;
 AC AAB93497;
 XX 26-JUN-2001 (first entry)
 DT 26-JUN-2001 (first entry)
 XX Human protein sequence SEQ ID NO:12811.
 DE Human protein sequence SEQ ID NO:12811.

KW Human; ribonucleotide reductase; cancer; DNA repair; p53.
 XX Homo sapiens.
 XX WO200100799-A1.
 XX 04-JAN-2001.
 XX 27-JUN-2000; 2000WO-JP004189.
 XX 28-JUN-1999; 99JP-00181131.
 PR 06-JUL-1999; 99JP-00192391.
 PR 21-JAN-2000; 2000JP-00017770.
 XX (TAKE) TAKEDA CHEM IND LTD.
 PA (NAKA/) NAKAMURA Y.
 XX Nakamura Y, Arakawa H, Tanaka H;
 PI WPI; 2001-112446/12.
 DR N-PSDB; AAF32438.
 XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
 PT diagnosis, treatment and prevention of cancer.
 XX Claim 1; Fig 1-3; 102pp; Japanese.
 XX The present sequence represents a human ribonucleotide reductase
 CC designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of
 CC the DNA repair mechanism and its activity is induced by p53. It can be
 CC used for the treatment, prevention and diagnosis of a wide range of
 CC cancers
 XX SQ Sequence 351 AA;
 Query Match 100.0%; Score 1821; DB 4; Length 351;
 Best Local Similarity 100.0%; Pred. No. 5.2e-179;
 Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGDPERPAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWKMYKQAO 60
 Db 1 MGDPERPAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWKMYKQAO 60
 QY 61 ASFWTAEVDLSKDLPHWNKLKADSKYFISHILAFPAASDGI VNNLVERFSQEVQVPEA 120
 Db 61 ASFWTAEVDLSKDLPHWNKLKADSKYFISHILAFPAASDGI VNNLVERFSQEVQVPEA 120
 QY 121 RCFYGFQIILIENVHSEMYSLIIDTYIRDPKKREFLFNAIETMPYVKKKADWALRWIADRK 180
 Db 121 RCFYGFQIILIENVHSEMYSLIIDTYIRDPKKREFLFNAIETMPYVKKKADWALRWIADRK 180
 QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDFACLMPQ 240
 Db 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDFACLMPQ 240
 QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVGLGFS 300
 Db 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVGLGFS 300
 QY 301 KVFQAEENPFDPMENISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF 351
 Db 301 KVFQAEENPFDPMENISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF 351
 RESULT 3
 AAB93497
 ID AAB93497 standard; protein; 351 AA.
 XX AAB93497;
 AC AAB93497;
 XX 26-JUN-2001 (first entry)
 DT 26-JUN-2001 (first entry)
 XX Human protein sequence SEQ ID NO:12811.
 DE Human protein sequence SEQ ID NO:12811.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX Homo sapiens.
 OS EP1074617-A2.
 PN 07-FEB-2001.
 PD 28-JUL-2000; 2000EP-00116126.
 XX 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX (HELI-) HELIX RES INST.
 PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.
 DR Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
 XX Claim 8; SEQ ID NO 12811; 2537bp + Sequence Listing; English.
 XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention
 XX Sequence 351 AA;
 SQ

Query Match 99.8%; Score 1817; DB 4; Length 351;
 Best Local Similarity 99.7%; Pred. No. 1.4e-178;
 Matches 350; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPPERAAAGLDQDERSSDTNESEIKNEBPLLRKSSRRFVIFPIQYDPIDWKMYKQAO 60
 DB 1 MGDPPERAAAGLDQDERSSDTNESEIKNEBPLLRKSSRRFVIFPIQYDPIDWKMYKQAO 60
 QY 61 ASFWTAEEVDLSKLPWNKLKADKEYFTSHILAPFAASDGIWNLVERFSGVOVPEA 120
 DB 61 ASFWTAEEVDLSKLPWNKLKADKEYFTSHILAPFAASDGIWNLVERFSGVOVPEA 120
 QY 121 RCFYGFQILINNVHSEMSLLIDTYRDPKREFLFNAITETMPYVKKKADWALRWIADRK 180
 DB 121 RCFYGFQILINNVHSEMSLLIDTYRDPKREFLFNAITETMPYVKKKADWALRWIADRK 180
 QY 181 STFGERVVAFVAVGVFSGSFAAIFWLKKGRLMPLGTFNSNELISRDEGLHCDFACLMPQ 240

DB 181 STFGERVVAFVAVGVFSGSFAAIFWLKKGRLMPLGTFNSNELISRDEGLHCDFACLMPQ 240
 QY 241 YLVNKPSEERREIIVDAVKIEQRFLEALPVGLIGNLCILMKQYIEFVADRLVLELGS 300
 DB 241 YLVNKPSEERREIIVDAVKIEQRFLEALPVGLIGNLCILMKQYIEFVADRLVLELGS 300
 QY 301 KVFQAEPPDFMENISLEGKTNFPEKRVSEYQRFVAVMAETDNDVFTLDADF 351
 DB 301 KVFQAEPPDFMENISLEGKTNFPEKRVSEYQRFVAVMAETDNDVFTLDADF 351

RESULT 4
 AAB69051
 ID AAB69051 standard; protein; 389 AA.
 XX AAB69051;
 AC 18-APR-2001 (first entry)
 DT Human R2 protein sequence SEQ ID NO:5.
 DE Human; ribonucleotide reductase; cancer; DNA repair; p53.
 XX Homo sapiens.
 OS WO200100799-A1.
 PN 04-JAN-2001.
 XX 27-JUN-2000; 2000WO-JP004189.
 XX 28-JUN-1999; 99JP-00181131.
 PR 06-JUL-1999; 99JP-00192391.
 PR 21-JAN-2000; 2000JP-00017770.
 XX (TAKE) TAKEDA CHEM IND LTD.
 PA (NAKA/) NAKAMURA Y.
 XX Nakamura Y, Arakawa H, Tanaka H;
 PI WPI; 2001-112446/12.
 XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for diagnosis, treatment and prevention of cancer.
 PT Example 2; Page 92-94; 102pp; Japanese.
 XX The present invention describes a human ribonucleotide reductase designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of the DNA repair mechanism and its activity is induced by p53. It can be used for the treatment, prevention and diagnosis of a wide range of cancers. The present sequence represents the human R2 protein which is used in an example from the present invention
 XX Sequence 389 AA;
 SQ

Query Match 79.8%; Score 1454; DB 4; Length 389;
 Best Local Similarity 83.8%; Pred. No. 4.8e-141;
 Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY 31 EEPPLRKSSRRFVIFPIQYDPIDWKMYKQAOASFTWAEVDLSKLPWNKLKADKEYFIS 90
 DB 69 DEPLLRENPRRFVIFPIEYHDIMQYKKAESFTWAEVDLSKLPWNKLKADKEYFIS 128
 QY 91 HILAPFAASDGIWNLVERFSGVOVPEARCFYGFQILINNVHSEMSLLIDTYRDPK 150
 DB 129 HVLAPFAASDGIWNLVERFSGVOVPEARCFYGFQILINNVHSEMSLLIDTYRDPK 188
 QY 151 KREFLFNAITETMPYVKKKADWALRWIADRKSTFGERVVAFVAVGVFSGSFAAIFWLK 210
 DB 189 EREFLFNAITETMPYVKKKADWALRWIADRKSTFGERVVAFVAVGVFSGSFAAIFWLK 248

QY 211 RGLMPLTFSNELISRDEGLHCDFACLMPQYLVNRPSEERVRRIIVDAVKIQBFLEAL 270
 Db 249 RGLMPLTFSNELISRDEGLHCDFACLMPKHLVHKPSEERVRRIIVAVRIQBFLEAL 308
 QY 271 PVGLIGMNCILMKQYIEFVADRLVLLVELGSKVFOAENPDPFNMENISLEGKTNFFKRVSE 330
 Db 309 PVKLGMMCTLMKQYIEFVADRLVLLVELGSKVFRVENPDPFNMENISLEGKTNFFKRVGE 368
 QY 331 YQRFVMAETTDNVFTLDADF 351
 Db 369 YQRMGMSSPTENSFTLDADF 389

RESULT 5
 ABU07433
 ID ABU07433 standard; protein; 389 AA.
 AC ABU07433;
 XX
 DT 28-JAN-2003 (first entry)
 XX
 DE Protein differentially regulated in prostate cancer #36.
 XX
 KW Prostate cancer; gene expression; differential regulation;
 KW molecular marker; drug target; cancer detection; cancer diagnosis;
 KW cancer staging; cancer grading; cancer assessing; cancer monitoring.
 XX
 OS Homo sapiens.
 PN WO200281638-A2.
 XX
 PD 17-OCT-2002.
 XX
 PF 08-APR-2002; 2002WO-US010824.
 XX
 PR 06-APR-2001; 2001US-0281731P.
 PR 06-APR-2001; 2001US-0281732P.
 XX
 PA (ORIG-) ORIGENE TECHNOLOGIES INC.
 XX
 PI Sun Z, Jay G;
 XX
 DR WPI; 2003-058520/05.
 DR N-PSDB; ABX10335.
 XX
 PT Novel genes which are differentially regulated in prostate cancer, useful
 PT for diagnosing prostate cancer in prostate tissue sample and assessing
 PT therapeutic or preventive intervention in prostate cancer patients.

Claim 1; Page 273-274; 416pp; English.

The invention describes genes (I) which are differentially regulated in prostate cancer. (I) Is useful for diagnosing a prostate cancer in a sample comprising prostate tissue, which involves determining the number of target genes which are differentially-regulated in the sample, where the number is indicative of the probability that the sample comprises prostate cancer. (I) is useful for assessing a therapeutic or preventive intervention in a subject having a prostate cancer, which involves determining the expression levels in a sample comprising prostate tissue of target genes which are differentially-regulated in prostate cancer. Preferably, the expression levels of at least 10 genes are determined. (I) is also useful for identifying agents that modulate a biological activity of a polypeptide differentially-regulated in prostate cancer cells, which involves contacting a polypeptide differentially-regulated in prostate cancer cells with a test agent under conditions effective for the test agent to modulate a biological activity of the polypeptide, and determining whether the test agent modulates the biological activity. (I) is useful as molecular markers, as drug targets, and for detecting, diagnosing, staging, grading, assessing, monitoring, prognosticating, preventing or treating, determining predisposition to diseases and conditions especially relating to prostate cancer. (I) and its expression products are used in the diagnostic test to assay for presence of cancer e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in

CC blood etc. (I) is useful for assessing cancer e.g., to determine the type
 CC of cancer, its stage of development, the nature of genetic defect, etc.
 CC The polypeptide encoded by (I) can be used as target for therapy or drug
 CC discovery. (I) can also be used for expressing the polypeptide and thus
 CC for searching specific binding partners of the polypeptide. (I) is useful
 CC in therapeutic applications to treat prostate cancer. The identification
 CC of specific genes, and groups of genes, expressed in pathways
 CC physiologically relevant to prostate cancer permits the definition of
 CC functional and disease pathways and the delineation of targets in these
 CC pathways which are useful in diagnostic, therapeutic, and clinical
 CC applications. This is the amino acid sequence of a protein differentially
 CC regulated in prostate cancer
 XX
 SQ Sequence 389 AA;

Query Match 79.8%; Score 1454; DB 6; Length 389;
 Best Local Similarity 83.8%; Pred. No. 4.8e-141;
 Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY 31 EEPLLRKSRFFVIFPIQYPDIMWKYKQAQASFWTAEEVDLSKDLPHWNKLKADKPYFIS 90
 Db 69 DEPLLRNPRFVIFPIEYHDIMWKYKAEASFWTAEEVDLSKDIQHWESLKPERYFIS 128
 QY 91 HILAPFAASDGI VNNLVERPSEVQVPRARCFYGFOLLENVHSEMYSLIDYIRDPK 150
 Db 129 HVLAPFAASDGI VNNLVERPSEVQVTEARCFYGFQIAMENIHSEMYSLIDYIKDPK 188
 QY 151 KREFLENAIETMPYKKKADWALRWIADRKSTGERVVAEAAVEGVFSGSFAAIFWLKK 210
 Db 189 EREFLENAIETMPCVKKKADWALRWIGDKEATYGERVVAFAAVEGIFSGSFASIFWLKK 248
 QY 211 RGLMPLTFSNELISRDEGLHCDFACLMPQYLVNKPSEERVRRIIVDAVKIQBFLEAL 270
 Db 249 RGLMPLTFSNELISRDEGLHCDFACLMPKHLVHKPSEERVRRIIVAVRIQBFLEAL 308
 QY 271 PVGLIGMNCILMKQYIEFVADRLVLLVELGSKVFOAENPDPFNMENISLEGKTNFFKRVSE 330
 Db 309 PVKLGMMCTLMKQYIEFVADRLVLLVELGSKVFRVENPDPFNMENISLEGKTNFFKRVGE 368
 QY 331 YQRFVMAETTDNVFTLDADF 351
 Db 369 YQRMGMSSPTENSFTLDADF 389

RESULT 6

ADJ68753
 ID ADJ68753 standard; protein; 389 AA.
 XX
 AC ADJ68753;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Human heat mitochondrial protein as a therapeutic target SeqID559.
 XX
 KW Mitochondrial; human; screening assay; diabetes mellitus;
 KW Huntington's disease; osteoarthritis;
 KW Leber's hereditary optic neuropathy; LHON;
 KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
 KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
 KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
 KW osteopathic; ophthalmological; cytostatic.
 XX
 OS Homo sapiens.
 XX
 PN WO2003087768-A2.
 XX
 PD 23-OCT-2003.
 XX
 PF 04-APR-2003; 2003WO-US010870.
 XX
 PR 12-APR-2002; 2002US-0372843P.
 PR 17-JUN-2002; 2002US-0389987P.
 PR 20-SEP-2002; 2002US-0412418P.

XX (MITO-) MITOKOR.
PA (BUCK-) BUCK INST AGE RES.
XX
XX
XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
PI Warnock DE;
XX
XX WPI; 2003-845369/78.
XX
XX Identifying a mitochondrial target for drug screening assays and for
PT treating diseases associated with altered mitochondrial function,
PT comprises detecting a modified polypeptide in a sample and correlating
PT with the disease.
XX
XX Claim 1; SEQ ID NO 559; 180pp; English.
XX
XX This invention relates to novel mitochondrial targets that can be used
CC for therapeutic intervention in treating a disease associated with
CC altered mitochondrial function. Specifically, it refers to a method for
CC identifying proteins of the human heart mitochondrial proteome that are
CC useful for drug screening assays, as well as therapeutic targets. The
CC present invention describes a method for identifying such proteins that
CC can be used in the treatment of various diseases associated with altered
CC mitochondrial function including diabetes mellitus, Huntington's disease,
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC compositions have neuroprotective, nootropic, antidiabetic,
CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
CC cytoskeletal activities. This polypeptide sequence is a human heart
CC mitochondrial protein of the invention.
XX
XX Sequence 389 AA;
XX
Query Match 79.8%; Score 1454; DB 7; Length 389;
Best Local Similarity 83.8%; Pred. No. 4.8e-141;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;
QY 31 EEPILKSSRRFVIFPIQYDPDIWKMYKQAQASFWTAEEVDLSKDLPHWNKLKADKCYFTS 90
DB 69 DEPLLRENPRFVIFPIEYHDIWQMYKKAASFWTAEEVDLSKDIQHWESLKPPEERYFTS 128
QY 91 HILAFPAASDGIWNLNVERFSQVQVPEARCFYGFQIILNHNHSEMYSLLDITYIRDPK 150
DB 129 HVLAFPAASDGIWNLNVERFSQVQVTEARCFYGFQIAMENIHSEMYSLLDITYIKDPK 188
QY 151 KREFLFNAIETMPYVKKKADWALRWIADRKSTFGERVVAFVAAVEGFFSGSFAAIFWLKK 210
DB 189 EREFLFNAIETMPCVKKKADWALRWIGDKKATYGERVVAFVAAVEGFFSGSPASIFWLKK 248
QY 211 RGLMPLGLTFSNELISRDEGLHCDPACLMFQYLVNKPSEVRREIIVDAVKIQEFTLEAL 270
DB 249 RGLMPLGLTFSNELISRDEGLHCDPACLMFQYLVNKPSEVRREIIVDAVKIQEFTLEAL 308
QY 271 PVGLIGMNCILMKQYIEFVADRLVLLVGLGSKYQFAENPFDMENISLEGKTNPFPEKRVGE 330
DB 309 PVKLLIGMNCILMKQYIEFVADRLVLLVGLGSKYQFVRNPFDMENISLEGKTNPFPEKRVGE 368
QY 331 YQFVAVMAETNDNVFTLDADF 351
DB 369 YQRMGMSSPTENSFTLDADF 389
RESULT 7
ID ADJ66564
XX ADJ66564 standard; protein; 389 AA.
XX
XX ADJ66564;
AC
XX
XX 06-MAY-2004 (first entry)
DT
XX
XX Ribonucleoside-diphosphate reductase M2 chain for anti-cancer complex.
DE
XX

KW neuroprotective; cytostatic; gene therapy; protein complex;
KW cellular network; cancer; neurodegenerative disease; drug target.
OS Homo sapiens.
PN WO2004009622-A2.
XX
XX 29-JAN-2004.
XX
XX 18-JUL-2003; 2003WO-EP007835.
XX
XX 19-JUL-2002; 2002EP-00016109.
PR 19-JUL-2002; 2002EP-00016111.
PR 19-JUL-2002; 2002EP-00016123.
PR 19-JUL-2002; 2002EP-00016128.
PR 22-JUL-2002; 2002EP-00016427.
XX
XX (CELL-) CELLZOME AG.
PA
XX Merino A, Bouwmeester T, Bauer A, Drewes G, Marzioch M, Kruse U;
PI Superti-Furga G, Eberhard D, Ruffner H, Hobson S, Helftenbein G;
PI Cruciat C;
XX
XX WPI; 2004-123372/12.
XX
XX New protein complexes of cellular networks underlying the development of
PT cancer and other diseases, useful for diagnosing and/or treating
PT neurodegenerative diseases or cancer, and in drug screening.
XX
XX Disclosure; SEQ ID NO 94; 809pp; English.
XX
XX The invention relates to a protein complex of cellular networks
CC underlying the development of cancer and other diseases. The complex (I)
CC comprises at least one first and second proteins selected from any of the
CC proteins listed in the specification, or their functionally active
CC derivatives, fragments, homologues or variants, the variants being
CC encoded by a nucleic acid that hybridizes to the nucleic acid encoding
CC the protein under low stringency conditions. A complex (II) comprises at
CC least two of the second proteins, where the low stringency conditions
CC comprise hybridization in a buffer comprising 35% formamide, 5 x SSC, 50
CC mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% BSA, 100 microg/ml
CC denatured salmon sperm DNA, and 10% (wt/vol) dextran sulfate for 18-20
CC hours at 40 deg C, washing in a buffer consisting of 2 x SSC, 25 mM Tris-
CC HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS for 1.5 hours at 55 deg C, and
CC washing in a buffer consisting of 2 x SSC, 25 mM Tris-HCl (pH 7.4), 5 mM
CC EDTA, and 0.1% SDS for 1.5 hours at 60 deg C. The composition and methods
CC are useful in diagnosing or treating diseases and disorders, preferably
CC neurodegenerative diseases. These may also be used as a drug target or in
CC manufacturing a medicament for the treatment or prevention of the above-
CC mentioned diseases or disorders. The composition may also be used for
CC treating cancer. This sequence represents one of the proteins of the
CC complex of the invention.
XX
XX Sequence 389 AA;
XX
Query Match 79.8%; Score 1454; DB 8; Length 389;
Best Local Similarity 83.8%; Pred. No. 4.8e-141;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;
QY 31 EEPILKSSRRFVIFPIQYDPDIWKMYKQAQASFWTAEEVDLSKDLPHWNKLKADKCYFTS 90
DB 69 DEPLLRENPRFVIFPIEYHDIWQMYKKAASFWTAEEVDLSKDIQHWESLKPPEERYFTS 128
QY 91 HILAFPAASDGIWNLNVERFSQVQVPEARCFYGFQIILNHNHSEMYSLLDITYIRDPK 150
DB 129 HVLAFPAASDGIWNLNVERFSQVQVTEARCFYGFQIAMENIHSEMYSLLDITYIKDPK 188
QY 151 KREFLFNAIETMPYVKKKADWALRWIADRKSTFGERVVAFVAAVEGFFSGSFAAIFWLKK 210
DB 189 EREFLFNAIETMPCVKKKADWALRWIGDKKATYGERVVAFVAAVEGFFSGSPASIFWLKK 248
QY 211 RGLMPLGLTFSNELISRDEGLHCDPACLMFQYLVNKPSEVRREIIVDAVKIQEFTLEAL 270

```

Db      249 RGLMPGLTFSNELISRDEGLHCDPACIMFKHLVHKPSEERVREIIINAVRIEQEFTEAL 308
QY      271 PVGLIGMNCILMKQYIEFVADRLVLVGLGSKVFOAENPFDPMENISLEGKTNPFKEKRVSE 330
Db      309 PVKLLGMNCTLMKQYIEFVADRLMLGLGSKVFRVENPFDPMENISLEGKTNPFKEKRVSE 368
QY      331 YORFAVMAETTDNVFTLDADF 351
Db      369 YORMGVMSSTENSFTLDADF 389

RESULT 8
ADK70435
ID      ADK70435 standard; protein; 389 AA.
XX
AC      ADK70435;
XX
DT      06-MAY-2004 (first entry)
XX
DE      Respiratory disease differentially expressed protein #1.
XX
KW      cytostatic; respiratory; antiasthmatic; gene therapy;
KW      differential gene expression; respiratory disorder; lung cancer;
KW      chronic obstructive pulmonary disease; emphysema; asthma.
XX
OS      Homo sapiens.
XX
PN      WO2003101283-A2.
XX
PD      11-DEC-2003.
XX
PF      02-JUN-2003; 2003WO-US017409.
XX
PR      04-JUN-2002; 2002US-0386005P.
XX
PA      (INCY-) INCYTE CORP.
XX
PI      Rickert PK, Krasnow R;
XX
PI      WPI; 2004-042945/04.
XX
DR
XX
XX      New combination comprising cDNAs and proteins that are differentially
XX      expressed in respiratory disorders, useful for diagnosing or treating
XX      respiratory diseases e.g. lung cancer, chronic obstructive pulmonary
XX      diseases or asthma.
XX
PS      Claim 14; SEQ ID NO 171; 343pp; English.
XX
CC      The invention relates to cDNA sequences that are differentially expressed
XX      in respiratory disorders or their complements or encoded proteins. The
XX      cDNAs and proteins are useful for diagnosing, treating or monitoring
XX      treatment of a subject with a respiratory disease including lung cancer,
XX      chronic obstructive pulmonary diseases, emphysema or asthma. The protein
XX      is also useful for screening molecules or compounds to identify at least
XX      one ligand which specifically binds the protein. It is also useful for
XX      preparing and purifying a polyclonal or monoclonal antibody. This
XX      sequence corresponds to a protein of the invention.
XX
SQ      Sequence 389 AA;

Query Match          79.8%; Score 1454; DB 8; Length 389;
Best Local Similarity 83.8%; Pred. No. 4.8e-141;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY      31 EEPLLRKSSRRVFIPFIQYPDIMWYKQAQASFMTAEVDLSKDLPHWNKLKADEKYFIS 90
Db      69 DEPLLRENPRFVFIPFIEYHDIMWYKKAESFMTAEVDLSKDIQHWESLKPERYFIS 128
QY      91 HILAFPAASDGI VNNLVRFSQEVQVPEARCFYGFQIILNVHSEMYSLLLIDTYIRDPK 150
Db      129 HVLAFPAASDGI VNNLVRFSQEVQITEARCFYGFQIAMENIHSEMYSLLLIDTYIKDPK 188
QY      151 KREFLFNAIETMPYVKKADWALRWIADRKSTFGERVVAFAAVEGFFSGSFAAIFWLKK 210

:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
189 EREFLNAIETMPYVKKADWALRWIADRKSTFGERVVAFAAVEGFFSGSFAAIFWLKK 248

:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
211 RGLMPGLTFSNELISRDEGLHCDPACIMFKHLVHKPSEERVREIIINAVRIEQEFTEAL 270

:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
249 RGLMPGLTFSNELISRDEGLHCDPACIMFKHLVHKPSEERVREIIINAVRIEQEFTEAL 308

:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
271 PVGLIGMNCILMKQYIEFVADRLVLVGLGSKVFOAENPFDPMENISLEGKTNPFKEKRVSE 330

:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
309 PVKLLGMNCTLMKQYIEFVADRLMLGLGSKVFRVENPFDPMENISLEGKTNPFKEKRVSE 368

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
331 YORFAVMAETTDNVFTLDADF 351

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
369 YORMGVMSSTENSFTLDADF 389

RESULT 9
ADN03789
ID      ADN03789 standard; protein; 389 AA.
XX
AC      ADN03789;
XX
DT      01-JUL-2004 (first entry)
XX
DE      Antipsoriatic protein sequence #91.
XX
KW      antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS      Homo sapiens.
XX
PN      WO2004028479-A2.
XX
PD      08-APR-2004.
XX
PF      25-SEP-2003; 2003WO-US030907.
XX
PR      25-SEP-2002; 2002US-0414006P.
XX
XX      (GETH ) GENENTECH INC.
XX
PI      Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI      Wu TD;
XX
XX      WPI; 2004-305105/28.
XX      N-PSDB; ADN03788.
XX
XX      New PRO nucleic acid or polypeptide, useful for preparing a
XX      pharmaceutical composition for diagnosing or treating psoriasis in a
XX      mammal.
XX
PS      Claim 9; SEQ ID NO 183; 3069pp; English.
XX
CC      The invention relates to novel polynucleotide and polypeptides for
XX      treating psoriasis or a sequence having at least 80% identity to the
XX      above sequences. The nucleic acid is useful for preparing a composition
XX      for diagnosing or treating psoriasis in a mammal. This sequence
XX      corresponds to one of the polypeptides of the invention.
XX
SQ      Sequence 389 AA;

Query Match          79.8%; Score 1454; DB 8; Length 389;
Best Local Similarity 83.8%; Pred. No. 4.8e-141;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY      31 EEPLLRKSSRRVFIPFIQYPDIMWYKQAQASFMTAEVDLSKDLPHWNKLKADEKYFIS 90
Db      69 DEPLLRENPRFVFIPFIEYHDIMWYKKAESFMTAEVDLSKDIQHWESLKPERYFIS 128
QY      91 HILAFPAASDGI VNNLVRFSQEVQVPEARCFYGFQIILNVHSEMYSLLLIDTYIRDPK 150
Db      129 HVLAFPAASDGI VNNLVRFSQEVQITEARCFYGFQIAMENIHSEMYSLLLIDTYIKDPK 188
QY      151 KREFLFNAIETMPYVKKADWALRWIADRKSTFGERVVAFAAVEGFFSGSFAAIFWLKK 210

```

Db 189 EREFLNAIETMPCVKKKADWALRWIGDKATYGERVVAFAAVEGIFSGSFASIFWLKK 248
Qy 211 RGLMPGLTFSNELISRDEGLHCDFACLMFOYLVNKPSEERVRRIIIVDAVKIEQEFLEAL 270
Db 249 RGLMPGLTFSNELISRDEGLHCDFACLMFKHLVHKPSEERVRRIIINAVRIQEFLEAL 308
Qy 271 PVGLIGNCMLMKQYIEFVADRLVDELGFSKVFOAENPFDPMENISLEGKTNFFEKRVSE 330
Db 309 PVKLGNCMLMKQYIEFVADRLVDELGFSKVFRVENPFDPMENISLEGKTNFFEKRVSE 368
Qy 331 YORFVAVMAETTDNVFTLDADF 351
Db 369 YQRMGMVMSPTENSFTLDADF 389

RESULT 10
ADN04444
ID ADN04444 standard; protein; 389 AA.
XX AC ADN04444;
XX DT 01-JUL-2004 (first entry)
XX DE Antipsoriatic protein sequence #415.
XX KW antipsoriatic; gene therapy; psoriasis; diagnosis.
XX OS Homo sapiens.
XX PN WO2004028479-A2.
XX PD 08-APR-2004.
XX PF 25-SEP-2003; 2003WO-US030907.
XX PR 25-SEP-2002; 2002US-0414006P.
XX PA (GETH) GENENTECH INC.
XX PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI Wu TD;
XX DR N-PSDB; ADN04443.
XX PT New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.
XX PS Claim 9; SEQ ID NO 838; 3069pp; English.
XX CC The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polypeptides of the invention.
XX SQ Sequence 389 AA;

Query Match 79.8%; Score 1454; DB 8; Length 389;
Best Local Similarity 83.8%; Pred. No. 4.8e-141;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;
Qy 31 EEPILKRSRFRVPIPIQVPOIKWYKQAQSFWTAEVDLSKDLPHWNKLKADKRYFS 90
Db 69 DEPLRENPRFVPIPIYHDIWQYKKAESFWTAEVDLSKDIQHWESLPKERYFIS 128
Qy 91 HILAFPAASDGVNENLVERFSQEQVPEARCFYGFQILINVHSEMYSLLDITVIRDPK 150
Db 129 HVLAFPAASDGVNENLVERFSQEQVTEARCFYGFQIAMENIHSEMYSLLDITVIKDPK 188
Qy 151 KREFLNAIETWPKKADWALRWIADRKSTFGSERVVAFAAVEGIFSGSFASIFWLKK 210

Db 189 EREFLNAIETMPCVKKKADWALRWIGDKATYGERVVAFAAVEGIFSGSFASIFWLKK 248
Qy 211 RGLMPGLTFSNELISRDEGLHCDFACLMFOYLVNKPSEERVRRIIIVDAVKIEQEFLEAL 270
Db 249 RGLMPGLTFSNELISRDEGLHCDFACLMFKHLVHKPSEERVRRIIINAVRIQEFLEAL 308
Qy 271 PVGLIGNCMLMKQYIEFVADRLVDELGFSKVFOAENPFDPMENISLEGKTNFFEKRVSE 330
Db 309 PVKLGNCMLMKQYIEFVADRLVDELGFSKVFRVENPFDPMENISLEGKTNFFEKRVSE 368
Qy 331 YORFVAVMAETTDNVFTLDADF 351
Db 369 YQRMGMVMSPTENSFTLDADF 389

RESULT 11
ADO19226
ID ADO19226 standard; protein; 389 AA.
XX AC ADO19226;
XX DT 12-AUG-2004 (first entry)
XX DE Human PRO polypeptide #79.
XX KW Human; PRO; immune related disorder; systemic lupus erythematosus;
KW rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis;
KW systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis;
KW autoimmune haemolytic anaemia; autoimmune thrombocytopenia; thyroiditis;
KW diabetes mellitus; renal disease; demyelinating disease;
KW central nervous system; peripheral nervous system;
KW demyelinating polynuropathy; Guillain-Barre syndrome;
KW chronic inflammatory demyelinating polynuropathy.
XX OS Homo sapiens.
XX PN WO2004043361-A2.
XX PD 27-MAY-2004.
XX PF 06-NOV-2003; 2003WO-US035268.
XX PR 08-NOV-2002; 2002US-0425235P.
XX PA (GETH) GENENTECH INC.
XX PI Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
PI Wood WI, Wu TD;
XX DR WPI; 2004-420067/39.
DR N-PSDB; ADO19225.
XX PT Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
PT treating an immune related disorder such as systemic lupus erythematosus,
PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
PT spondyloarthopathy.
XX PS Claim 7; SEQ ID NO 158; 1731pp; English.
XX CC The invention relates to human PRO polypeptides and the polynucleotides
CC encoding them. The polypeptides and polynucleotides are useful for
CC treating and diagnosing immune related disorders in mammals. The immune
CC related disorders include systemic lupus erythematosus, rheumatoid
CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic
CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
CC mellitus, immune-mediated renal disease, demyelinating diseases of the
CC central or peripheral nervous system, demyelinating polynuropathy,
CC Guillain-Barre syndrome and chronic inflammatory demyelinating
CC polynuropathy. This sequence represents a human PRO polypeptide of the
CC invention.

```
SQ Sequence 389 AA;
Query Match 79.8%; Score 1454; DB 8; Length 389;
Best Local Similarity 83.8%; Pred. No. 4.8e-141;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;
QY 31 EEPRLKSRFFVIFPIQYDPDIWKMVYKQAOASFWTAEEVDLSKDLPHWNKLADEKYFIS 90
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 69 DEPLLRENPRFVIFPIEYHDIMQMYKKAESFWTAEEVDLSKDIQHWESLKPEERYFIS 128
QY 91 HILAFPAASDGIWNLVERFSQEVQVPEARCFYGFQILLENVHSEMYSLIDTYIRDPK 150
Db 129 HVLAFPAASDGIWNLVERFSQEVQVPEARCFYGFQILLENVHSEMYSLIDTYIRDPK 188
QY 151 KREFLNAIETMPVYKKADWALRWIADRKSTFCGRVVAFAAEGVFSSGFAAIFWLKK 210
Db 189 EREFLNAIETMPVYKKADWALRWIADRKSTFCGRVVAFAAEGVFSSGFAAIFWLKK 248
QY 211 RGLMPGLTFSNELISRDEGLHCDPACLMFQYLVNKPSEERVREIIVDAVKIQEFLTEAL 270
Db 249 RGLMPGLTFSNELISRDEGLHCDPACLMFQYLVNKPSEERVREIIVDAVKIQEFLTEAL 308
QY 271 PVGLIGMNCILMKQYIEFVADRLLELVEGFSKVFQAEHPDFMENISLGKTNFFPKRVSE 330
Db 309 PVKLGIMNCTLMKQYIEFVADRLLELVEGFSKVFQAEHPDFMENISLGKTNFFPKRVSE 368
QY 331 YQRFVAVMAETTDNVTLDAF 351
Db 369 YQRMGMSSPTENSFTLDAF 389

RESULT 12
ADQ09272
ID ADQ09272 standard; protein; 389 AA.
XX
AC ADQ09272;
XX
DT 23-SEP-2004 (first entry)
XX
DE Human RRM2 protein SEQ ID NO:457.
XX
KW thanatos-associated protein; THAP; THAP responsive gene; THAP family;
KW THAP responsive element; angiogenesis; inflammation; metastasis; cancer;
KW apoptosis; cardiovascular disease; neurodegenerative disease; chemokine;
KW antiangiogenic; antiinflammatory; cardiovascular; cytostatic;
KW neuroprotective; osteopathic; THAP modulator; THAP synthesis modulator;
KW human.
XX
OS Homo sapiens.
XX
PN WO2004055050-A2.
XX
PD 01-JUL-2004.
XX
PF 10-DEC-2003; 2003WO-IB006434.
XX
PR 10-DEC-2002; 2002US-0432699P.
XX
PR 03-JUL-2003; 2003US-0485027P.
XX
PA (ENDO-) ENDOCUBE SAS.
XX
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI Girard J, Amalric F, Rousseigne M, Clouaire T;
XX
DR WPI; 2004-525034/50.
XX
DR N-PSDB; ADQ09273.
XX
Modulating expression of a Thanatos (death)-Associated Protein (THAP)
PT responsive gene for preventing or treating e.g. cancer or inflammation,
PT comprises modulating the interaction of a THAP polypeptide with a nucleic
PT acid.
XX
Example 47; SEQ ID NO 457; 612pp; English.
XX
PS
```

```
XX
CC The present invention describes a method for modulating the expression of
CC a thanatos (death)-associated protein (THAP) responsive gene. The method
CC comprises modulating the interaction of a THAP-family polypeptide or its
CC biological fragment with a nucleic acid, and so enhancing or repressing
CC the expression of the THAP responsive gene. Also described: (1) a method
CC of modulating the expression of a gene responsive to a THAP/chemokine
CC complex; (2) a pharmaceutical composition comprising a THAP responsive
CC element in a pharmaceutical carrier; (3) a transcription factor decoy
CC consisting essentially of a THAP responsive element; (4) a cell
CC comprising a transcription factor decoy described above; (5) methods of
CC modulating the interaction between a nucleic acid and a THAP-family
CC polypeptide or its biological fragment, or a nucleic acid and a
CC THAP/chemokine complex; (6) a vector packaging cell line comprising a
CC cell comprising a viral vector which comprises a promoter operably linked
CC to a nucleic acid encoding a THAP-family polypeptide or its biological
CC fragment; (7) a method of constructing a cell which expresses a
CC recombinant THAP-family polypeptide; (8) a method of ameliorating
CC symptoms associated with a condition mediated by a THAP/chemokine complex
CC ; (9) methods of identifying a test compound that modulates transcription
CC at a THAP responsive element or that modulates the transport of a
CC chemokine into the nucleus; (10) methods for reducing the symptoms
CC associated with a condition selected from excessive or insufficient
CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
CC or insufficient apoptosis, cardiovascular disease and neurodegenerative
CC diseases; symptoms associated with a condition resulting from the
CC activity of a chemokine or a THAP-family polypeptide in an individual; or
CC symptoms associated with transcriptional repression or activation
CC mediated by a THAP-family polypeptide in an individual; (11) a vector
CC encoding a THAP responsive promoter operably linked to a nucleic acid
CC comprising a detectable product; (12) a genetically engineered cell
CC comprising the vector described above or that expresses a THAP-family
CC polypeptide or its biological fragment; (13) an in vitro transcription
CC reaction comprising a nucleic acid comprising a THAP responsive promoter,
CC ribonucleotides and an RNA polymerase; and (14) an isolated mutant THAP-
CC family polypeptide that does not bind to a chemokine. The pharmaceutical
CC composition has antiangiogenic, antiinflammatory, cardiovascular,
CC cytostatic, neuroprotective and osteopathic activities, and can be used
CC as a THAP and THAP synthesis modulator. The composition can be used for
CC modulating the expression of a THAP responsive gene. Modulation is useful
CC for reducing symptoms of conditions such as excessive or insufficient
CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
CC or insufficient apoptosis, cardiovascular disease or neurodegenerative
CC diseases. The present sequence is used in the exemplification of the
CC present invention.
XX
SQ Sequence 389 AA;
```

```
Query Match 79.8%; Score 1454; DB 8; Length 389;
Best Local Similarity 83.8%; Pred. No. 4.8e-141;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;
```

```
QY 31 EEPRLKSRFFVIFPIQYDPDIWKMVYKQAOASFWTAEEVDLSKDLPHWNKLADEKYFIS 90
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 69 DEPLLRENPRFVIFPIEYHDIMQMYKKAESFWTAEEVDLSKDIQHWESLKPEERYFIS 128
QY 91 HILAFPAASDGIWNLVERFSQEVQVPEARCFYGFQILLENVHSEMYSLIDTYIRDPK 150
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 129 HVLAFPAASDGIWNLVERFSQEVQVTEARCFYGFQIAMENIHSEMYSLIDTYIKDPK 188
QY 151 KREFLNAIETMPVYKKADWALRWIADRKSTFCGRVVAFAAEGVFSSGFAAIFWLKK 210
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 189 EREFLNAIETMPVYKKADWALRWIADRKSTFCGRVVAFAAEGVFSSGFAAIFWLKK 248
QY 211 RGLMPGLTFSNELISRDEGLHCDPACLMFQYLVNKPSEERVREIIVDAVKIQEFLTEAL 270
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 249 RGLMPGLTFSNELISRDEGLHCDPACLMFQYLVNKPSEERVREIIVDAVKIQEFLTEAL 308
QY 271 PVGLIGMNCILMKQYIEFVADRLLELVEGFSKVFQAEHPDFMENISLGKTNFFPKRVSE 330
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 309 PVKLGIMNCTLMKQYIEFVADRLLELVEGFSKVFQAEHPDFMENISLGKTNFFPKRVSE 368
QY 331 YQRFVAVMAETTDNVTLDAF 351
```


CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC neotropic; vasotropic; antipsoriatic and antiangiogenic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention

XX SQ Sequence 413 AA;

Query Match 79.8%; Score 1454; DB 3; Length 413;
Best Local Similarity 83.8%; Pred. No. 5.2e-141;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY 31 EEPFLKSSRRFVIFPIQYDINWYKQAQASFTAEVDLSKDLPHWNKLKADKRYFIS 90
DB 93 DEPLLRENPRFRVIFPIEYHDINWYKKAASFTAEVDLSKDIQHWESLKPEERYFIS 152
QY 91 HILAPFAASDGINVENLVERFSQVQPEARCFYGFQILLENVHSEMYSLIDITYIDPK 150
DB 153 HVLAPFAASDGINVENLVERFSQVQPEARCFYGFQIAMENHSEMYSLIDITYIDPK 212
QY 151 KREFLFNAITMPYVKKKADWALRWIGDKATYGERVVAFAAEGVFFSGSFAIFWLKK 210
DB 213 EREFLNAITMPYVKKKADWALRWIGDKATYGERVVAFAAEGVFFSGSFAIFWLKK 272
QY 211 RGLMPGLTFSNELISRDEGLHCDPACLMFOYLNKPSSEVRVRIIIVDAVKIEQFTEAL 270
DB 273 RGLMPGLTFSNELISRDEGLHCDPACLMFKHLVHKPSEVRVRIIINAVRIEQLTEAL 332
QY 271 PVGLIGNCMLMKQYIEFVADRLLELGFSGKVPQAEENPFDMENISLEGTNFEKRVSE 330
DB 333 PVKLGIMNCTLMKQYIEFVADRLLELGFSGKVPQAEENPFDMENISLEGTNFEKRVSE 392
QY 331 YQRPAYMAETTDNVFTLDADF 351
DB 393 YQRMGMVSSPTENSFTLDADF 413

RESULT 15

AAU28017
ID AAU28017 standard; protein; 453 AA.

XX AAU28017;
AC AAU28017;
XX 18-DEC-2001 (first entry)

DE Human contig polypeptide sequence #170.

XX Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
KW nervous system disorder; inflammatory disorder; cell differentiation;
KW angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;
KW cytostatic; antirheumatic; antiarthritic; vulnetary; antinflammatory;
KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
KW neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;
KW immunostimulant; analgesic; gene therapy.

XX Homo sapiens.
OS Synthetic.
XX WO200164834-A2.

XX 07-SBP-2001.
XX 26-FEB-2001; 2001WO-US004926.
XX 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
PR 17-JUN-2000; 2000US-00597707.
PR 14-JUL-2000; 2000US-00616807.
PR 13-SEP-2000; 2000US-00664641.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;
PI Dmanac R;
XX WPI; 2001-589862/66.
DR N-PSDB; AAS44917.
XX Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis, treatment of cancer,
PT neurological, inflammatory disorders and for use in arrays for detection.
XX Claim 10; Page 146-147; 153pp; English.

CC Sequences AAU27676-AAU28019 represent full-length polypeptides and contig
CC polypeptides of the invention. The proteins and their associated DNA
CC sequences are useful for the treatment, diagnosis and prevention of
CC various types of disorder in a mammalian subject such as a human, dog,
CC monkey, mouse, hamster or rat. The disorders include cancers such as
CC leukaemia, lymphoma and neuroblastoma, autoimmune disorders such as
CC multiple sclerosis, connective tissue disease, rheumatoid arthritis,
CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's
CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
CC bowel disease. The sequences exhibit activity relating to angiogenesis,
CC cell proliferation, cell differentiation, stem cell growth factor,
CC activin or inhibin. Therefore, they can be used to manipulate stem cells
CC in culture to give rise to neuroepithelial cells that can be used to
CC augment or replace cells damaged by illness, accidental damage or genetic
CC disorders. The sequences may also be used for regeneration of bone,
CC cartilage, tendons and ligaments and in tissue repair and burn healing.
CC Note: Some sequences for this patent did not form part of the printed
CC specification, but were obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 453 AA;

Query Match 79.8%; Score 1454; DB 4; Length 453;
Best Local Similarity 83.8%; Pred. No. 6e-141;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY 31 EEPFLKSSRRFVIFPIQYDINWYKQAQASFTAEVDLSKDLPHWNKLKADKRYFIS 90
DB 133 DEPLLRENPRFRVIFPIEYHDINWYKKAASFTAEVDLSKDIQHWESLKPEERYFIS 192
QY 91 HILAPFAASDGINVENLVERFSQVQPEARCFYGFQILLENVHSEMYSLIDITYIDPK 150
DB 193 HVLAPFAASDGINVENLVERFSQVQPEARCFYGFQIAMENHSEMYSLIDITYIDPK 252
QY 151 KREFLFNAITMPYVKKKADWALRWIGDKATYGERVVAFAAEGVFFSGSFAIFWLKK 210
DB 253 EREFLNAITMPYVKKKADWALRWIGDKATYGERVVAFAAEGVFFSGSFAIFWLKK 312
QY 211 RGLMPGLTFSNELISRDEGLHCDPACLMFOYLNKPSSEVRVRIIIVDAVKIEQFTEAL 270
DB 313 RGLMPGLTFSNELISRDEGLHCDPACLMFKHLVHKPSEVRVRIIINAVRIEQLTEAL 372
QY 271 PVGLIGNCMLMKQYIEFVADRLLELGFSGKVPQAEENPFDMENISLEGTNFEKRVSE 330

Db 373 PVKLGMMCTLMKQYIEFVADRLMLELGFSGKVRVENPFDPMENISLEGKTNFFEKRVGE 432
Qy 331 YORFAVMAETTDNVFTLDADF 351
Db 433 YORMGVMSPTENSFTLDADF 453

Search completed: October 30, 2005, 06:56:17
Job time : 75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2005, 06:48:11 ; Search time 21 seconds
(without alignments)
1608.194 Million cell updates/sec

Title: US-10-698-228-1
Perfect score: 1821
Sequence: 1 MGDPERPEAGLDQDERSSS.....QRFVAVAEITDNTVTLADAF 351

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	100.0	366	2 T46249	hypothetical prote
2	1454	79.8	389	2 S25854	ribonucleoside-dip
3	1443	79.2	390	1 S06735	ribonucleoside-dip
4	1421	78.0	386	2 S27153	ribonucleoside-dip
5	1331	73.1	384	2 S24585	ribonucleoside-dip
6	1280	70.3	319	2 T30782	ribonucleoside-dip
7	1280	70.3	319	2 B72154	B4L protein - vari
8	1279	70.2	333	1 RDVZVV	ribonucleoside-dip
9	1279	70.2	333	2 T28466	ribonucleoside-dip
10	1266	69.5	333	2 H36839	C8L protein - vari
11	1258.5	69.1	348	2 T30470	ribonucleoside-dip
12	1217	66.8	299	1 RDSS2R	ribonucleoside-dip
13	1214	66.7	381	2 T19876	hypothetical prote
14	1168.5	64.2	399	1 A26916	ribonucleoside-dip
15	1161.5	63.8	391	2 T39992	ribonucleoside-dip
16	1161.5	63.8	391	2 S34808	ribonucleoside-dip
17	1128.5	62.0	329	2 T03688	ribonucleoside-dip
18	1070.5	58.8	340	2 S68538	ribonucleoside-dip
19	1024.5	56.3	349	2 B49412	ribonucleoside-dip
20	950	52.2	324	2 T17978	probable ribonucle
21	805	44.2	345	2 S59744	ribonucleoside-dip
22	681	37.4	668	2 T29884	hypothetical prote
23	638	35.0	327	1 RDVZAS	ribonucleoside-dip
24	385	21.1	331	2 A84389	ribonucleoside red
25	384.5	21.1	415	2 A83502	ribonucleoside red
26	358.5	19.7	346	2 G81728	ribonucleoside-dip
27	357.5	19.6	346	2 E71466	probable ribonucle
28	335	18.4	346	2 B72010	ribonucleoside-dip
29	335	18.4	346	2 F86613	ribonucleoside red

ALIGNMENTS

RESULT 1

T46249

hypothetical protein DKFZp761E1312.1 - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004

R;Anson, W.; Winkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, January 2000

A;Reference number: Z23028

A;Accession: T46249

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-366 <AAA>

A;Cross-references: UNIPROT:Q9NTD8; EMBL:AL137348

A;Experimental source: adult amygdala; clone DKFZp761E1312

C;Genetics:

A;Note: DKFZp761E1312.1

C;Superfamily: ribonucleoside reductase small subunit

Query Match 100.0%; Score 1821; DB 2; Length 366;
Best Local Similarity 100.0%; Pred. No. 1.6e-141;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGDPPEAAGLDQDERSSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYPIWKMYKQAQ	60
Db	16	MGDPPEAAGLDQDERSSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYPIWKMYKQAQ	75
Qy	61	ASFTAAEVDLSKDLPHWNKLGKADKDYFISHILAFFAASDGIIVNENLVERFSQEVQVPEA	120
Db	76	ASFTAAEVDLSKDLPHWNKLGKADKDYFISHILAFFAASDGIIVNENLVERFSQEVQVPEA	135
Qy	121	RCFYGQILLENVHSEMYSLIDTYIRDPKKREFLNATETMPYVKKADWALRIADRK	180
Db	136	RCFYGQILLENVHSEMYSLIDTYIRDPKKREFLNATETMPYVKKADWALRIADRK	195
Qy	181	STFGERVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDFACLMFQ	240
Db	196	STFGERVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDFACLMFQ	255
Qy	241	YLKNSSEERVRRIIVDAVKIEQFLTEALPVGLIGNLCILMKQYIEFVADRLVELGFS	300
Db	256	YLKNSSEERVRRIIVDAVKIEQFLTEALPVGLIGNLCILMKQYIEFVADRLVELGFS	315
Qy	301	KVFAENPFDFMENISLEGKTNFEKRVSEYQRFVAVNAETDNTVTLADAF	351
Db	316	KVFAENPFDFMENISLEGKTNFEKRVSEYQRFVAVNAETDNTVTLADAF	366

RESULT 2

S25854

ribonucleoside-diphosphate reductase (EC 1.1.7.4.1) small chain - human

N;Alternate names: ribonucleotide reductase M2 chain; ribonucleotide reductase small cha


```
QY 200 GSPAAIFWLKRGIMPLGLTFSNELISRDEGLHCDPACLMFQYLVNKPSEBVRVREIIVDAV 259
Db 239 GSPASIFWLKRGIMPLGLTFSNELISRDEGLHCDPACLMFQYLVNKPSEBVRVREIIVDAV 298
QY 260 KIQEELTEALPVGLIGMNCILMKQYIEFVADRLVLVGLGFSKVFOAENPDPFENISLEG 319
Db 299 RIEQEFTEALPVGLIGMNCILMKQYIEFVADRLVLVGLGFSKVFOAENPDPFENISLEG 358
QY 320 KTNFFPKRYSEYQRFAPVMAETTDNVFTLDADF 351
Db 359 KTNFFPKRYSEYQRFAPVMAETTDNVFTLDADF 386

RESULT 5
S24585
ribonucleoside-diphosphate reductase (EC 1.17.4.1) chain M2 - Atlantic surf clam
N;Alternate names: ibonucleotide-diphosphate reductase small chain
C;Species: Spisula solidissima (Atlantic surf clam)
C;Date: 22-Nov-1999 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S24585
R;Hunt, T.; Standart, N.
submitted to the EMBL Data Library, November 1990
A;Reference number: S24584
A;Accession: S24585
A;Molecule type: mRNA
A;Residues: 1-384 <HUN>
A;Cross-references: UNIPROT:P07201; EMBL:X55125; NID:g10341; PIDN:CAA38919.1; PID:g10342
C;Superfamily: ribonucleoside reductase small subunit
C;Keywords: oxidoreductase

Query Match 73.1%; Score 1331; DB 2; Length 384;
Best Local Similarity 74.3%; Pred. No. 2.4e-101; Mismatches 41; Indels 10; Gaps 4;
Matches 257; Conservative 38;

QY 16 ERSSTDN-ESEIK-----SNEEPLRKSRFRVFPQYDWMYKQAQAFWTAEE 68
Db 39 QRSTQKTLKQEIKPVKVKSQQVEPLADNPRFVLPQYHDIMWKYKKAESFWTAEE 98
QY 69 VDLKDLPHNKLKADEKVFISHILAFFAASDGIIVNENVERFSQVQPEARCFYGFQI 128
Db 99 VDLKDMAHWSLKEKEHFISHVLFAAASDGIIVNENVERFSQVQPEARCFYGFQI 158
QY 129 LIENHSEMYSLIDTYIRDPKRELENAETMPYVKKKADWALRWIADRKSTFGERVV 188
Db 159 AMENIHSEMYSLIDTYIRDPKRELENAETMPYVKKKADWALRWIADRKSTFGERVV 218
QY 189 AFAAEGVFFSGSFAAIFWLKRGIMPLGLTFSNELISRDEGLHCDPACLMFQYLVNKPSE 248
Db 219 AFAAEGVFFSGSFAAIFWLKRGIMPLGLTFSNELISRDEGLHCDPACLMFQYLVNKPSE 278
QY 249 ERVREIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLVLVGLGFSKVFOAENP 308
Db 279 ERHQIIDEAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLVLVGLGFSKVFOAENP 338
QY 309 FQFENISLEGKTNFFPKRYSEYQRFAPVMA--ETTD-NVFTLDADF 351
Db 339 FQFMEHISLEGKTNFFPKRYSEYQRFAPVMA--ETTD-NVFTLDADF 384

RESULT 6
T30782
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - vaccinia virus (strain
N;Alternate names: ribonucleotide reductase, small subunit
C;Species: vaccinia virus
A;Variety: strain Ankara
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T30782
R;Antoine, G.; Scheiflinger, F.; Falkner, F.G.; Dorner, F.
submitted to the EMBL Data Library, March 1997
A;Description: The complete genomic sequence of the Modified Vaccinia Ankara (MVA) strain
A;Reference number: Z20877
A;Accession: T30782
A;Status: preliminary; translated from GB/EMBL/DBJ
```

```
A;Molecule type: DNA
A;Residues: 1-319 <ANT>
A;Cross-references: UNIPROT:O57175; EMBL:U94848; PIDN:AAB96415.1
A;Experimental source: strain Ankara
C;Genetics:
A;Note: MVA032L
C;Function:
A;Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleoside
C;Superfamily: ribonucleoside reductase small subunit
C;Keywords: deoxyribonucleotide biosynthesis; oxidoreductase

Query Match 70.3%; Score 1280; DB 2; Length 319;
Best Local Similarity 74.7%; Pred. No. 2.8e-97;
Matches 239; Conservative 34; Mismatches 45; Indels 2; Gaps 2;

QY 32 EPLLKSRFRVFPQYDWMYKQAQAFWTAEEVDLSKDLPHNKLKADEKYPISH 91
Db 2 EPILANPRFVFPQYHDIMWKYKKAESFWTVEVDISKDINDNKLTPDEKYPFKH 61
QY 92 ILAFFAASDGIIVNENVERFSQVQPEARCFYGFQILNENHSEMYSLIDTYIRDPKK 151
Db 62 VLAFPAASDGIIVNENLAERFCTEQITEARCFYGFQWAIENIHSEMYSLIDTYVKSNE 121
QY 152 REFLFNAIETMPYVKKKADWALRWIADRKSTFGERVVFAAEGVFFSGSFAAIFWLKCR 211
Db 122 KNYLFNAIETMPYVKKKADWAKWIHD-SAGYGERLIAFAAEGVFFSGSFAAIFWLKCR 180
QY 212 GLMPGLTFSNELISRDEGLHCDPACLMFQYLVNKPSEBVRVREIIVDAVKIEQEFTEALP 271
Db 181 GLMPGLTFSNELISRDEGLHCDPACLMFQYLVNKPSEBVRVREIIVDAVKIEQEFTEALP 240
QY 272 VGLIGMNCILMKQYIEFVADRLVLVGLGFSKVFOAENPDPFENISLEGKTNFFPKRYSEY 331
Db 241 VKLIGMNCILMKQYIEFVADRLVLVGLGFSKVFOAENPDPFENISLEGKTNFFPKRYSEY 300
QY 332 QRPAVMAETTDNVFTLDADF 351
Db 301 QRKGVMSQ-EDNHFSLDADF 319

RESULT 7
B72154
E4L protein - variola minor virus (strain Garcia-1966)
C;Species: variola minor virus
C;Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 09-Jul-2004
C;Accession: B72154
R;Shchelkunov, S.N.; Totmenin, A.V.; Gutorov, V.V.; Safronov, P.P.; Massung, R.F.; Lopar
submitted to GenBank, March 1998
A;Description: Analysis of the complete coding sequence of DNA of alastrim variola minor
A;Reference number: A72150
A;Accession: B72154
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-319 <SHC>
A;Cross-references: UNIPROT:Q89087; GB:Y16780; NID:g5830555; PIDN:CAB54628.1; PID:g5830555
A;Experimental source: strain Garcia-1966
C;Genetics:
A;Gene: E4L
C;Superfamily: ribonucleoside reductase small subunit

Query Match 70.3%; Score 1280; DB 2; Length 319;
Best Local Similarity 75.0%; Pred. No. 2.8e-97;
Matches 240; Conservative 33; Mismatches 45; Indels 2; Gaps 2;

QY 32 EPLLKSRFRVFPQYDWMYKQAQAFWTAEEVDLSKDLPHNKLKADEKYPISH 91
Db 2 EPILANPRFVFPQYHDIMWKYKKAESFWTVEVDISKDINDNKLTPDEKYPFKH 61
QY 92 ILAFFAASDGIIVNENVERFSQVQPEARCFYGFQILNENHSEMYSLIDTYIRDPKK 151
Db 62 VLAFPAASDGIIVNENLAERFCTEQITEARCFYGFQWAIENIHSEMYSLIDTYVKSNE 121
QY 152 REFLFNAIETMPYVKKKADWALRWIADRKSTFGERVVFAAEGVFFSGSFAAIFWLKCR 211
```

Db 122 KNYLFNAIETMPCKVKADWAQKWIHD-SAGYGERLIAFAAAGVIFGSGFASIFWLKKR 180
QY 212 GLMPGLTFSNELSRDEGLHCDFACLMFQYLVNKPSEERVREIIVDAVKIEQEFLEALP 271
Db 181 GLMPGLTFSNELSRDEGLHCDFACLMFQYLVNKPSEERVREIIVDAVKIEQEFLEALP 240
QY 272 VGLIGNCMLMKQYIEFVADRLVELGFSKVOAENPFDPMENISLEGKTNFFEKRVSEY 331
Db 241 VKLIGNCENMKTYIEFVADRLVELGFSKVOAENPFDPMENISLEGKTNFFEKRVSEY 300
QY 332 QRFVMAETTDNVFTLDADF 351
Db 301 QKGMVMSQ-EDNHFSLDVDF 319
RESULT 8
RDVZWV
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - vaccinia virus
N:Alternate names: F4L protein
C:Species: vaccinia virus
C>Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
C:Accession: A29892; I42506; F36213
R:Slabaugh, M.; Roseman, N.; Davis, R.; Mathews, C.
J. Virol. 62, 519-527, 1988
A>Title: Vaccinia virus-encoded ribonucleotide reductase: sequence conservation of the
A:Reference number: A29892; MUID:88091062; PMID:2826813
A:Accession: A29892
A:Molecule type: DNA
A:Residues: 1-319 <SLA>
A:Cross-references: UNIPROT:P11159; GB:M19117; NID:G335808; PIDN:AAA88680.1; PID:G335809
A:Experimental source: strain WR
R:Goebel, S.J.; Johnson, G.P.; Perkus, M.E.; Davis, S.W.; Winslow, J.P.; Paoletti, E.
Virology 179, 517-563, 1990
A>Title: Appendix to "The complete DNA sequence of vaccinia virus".
A:Reference number: A42501
A:Accession: I42506
A:Molecule type: DNA
A:Residues: 1-212, 'Y', 214-319 <GOE>
A:Cross-references: GB:M35027; NID:G335317; PIDN:AAA48018.1; PID:G335366
A:Experimental source: strain Copenhagen
R:Goebel, S.J.; Johnson, G.P.; Perkus, M.E.; Davis, S.W.; Winslow, J.P.; Paoletti, E.
Virology 179, 247-266, 1990
A>Title: The complete DNA sequence of vaccinia virus.
A:Reference number: A42531; MUID:91021027; PMID:2219722
A:Contents: annotation; possible protein-coding frames
A>Note: neither amino acid nor nucleotide sequence is given
R:Roseman, N.A.; Slabaugh, M.B.
Virology 178, 410-418, 1990
A>Title: The vaccinia virus HindIII F fragment: nucleotide sequence of the left 6.2 kb.
A:Reference number: A36213; MUID:91020979; PMID:2219701
A:Accession: F36213
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-319 <ROS>
A:Cross-references: EMBL:M34358; NID:G335618; PIDN:AAA48244.1; PID:G335624
A:Experimental source: strain WR
C:Function:
A:Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleoside
C:Superfamily: ribonucleoside reductase small subunit
C:Keywords: DNA replication; iron; metalloprotein; oxidoreductase
F:70,101,104,163,197,200/Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, His) #sta
F:108/Active site: Tyr (stable tyrosyl radical) #status predicted
Query Match 70.2%; Score 1279; DB 1; Length 319;
Best Local Similarity 74.7%; Pred. No. 3.3e-97;
Matches 239; Conservative 34; Mismatches 45; Indels 2; Gaps 2;
QY 32 EPLLKRSRRFVIFPIQYDINWKYKQAQSFMTAEVDLSKDLPHNKLKADKYPFISH 91
Db 2 EPLIAPNPNRFVIFPIQYDINWKYKKAASFTVVEVDISKDINDWNKLTPEKYPFIKH 61
QY 92 ILAFAASDGINVENLVERSQEVOVPEARCFYGFQILLIENVHSEMYSLLLDITVYRDPKK 151

Db 62 VLAFFAASDGINVENLAERFCTEQITEARCFYGFQMAIENHSEMYSLLLDITVYKDSNE 121
QY 152 REFLFNAIETMPYKKADWALRIADRKSTFGFGRVVAFAAAGVIFGSGFASIFWLKKR 211
Db 122 KNYLFNAIETMPCKVKADWAQKWIHD-SAGYGERLIAFAAAGVIFGSGFASIFWLKKR 180
QY 212 GLMPGLTFSNELSRDEGLHCDFACLMFQYLVNKPSEERVREIIVDAVKIEQEFLEALP 271
Db 181 GLMPGLTFSNELSRDEGLHCDFACLMFQYLVNKPSEERVREIIVDAVKIEQEFLEALP 240
QY 272 VGLIGNCMLMKQYIEFVADRLVELGFSKVOAENPFDPMENISLEGKTNFFEKRVSEY 331
Db 241 VKLIGNCENMKTYIEFVADRLVELGFSKVOAENPFDPMENISLEGKTNFFEKRVSEY 300
QY 332 QRFVMAETTDNVFTLDADF 351
Db 301 QKGMVMSQ-EDNHFSLDVDF 319
RESULT 9
T28466
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - variola major virus
C:Species: variola major virus
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 15-Sep-2003
C:Accession: T28466
R:Massung, R.F.; Esposito, J.J.; Liu, L.I.; Qi, J.; Utterback, T.R.; Knight, J.C.; Aubin
Nature 366, 748-751, 1993
A>Title: Potential virulence determinants in terminal regions of variola smallpox virus
A:Reference number: Z20488; MUID:94088747; PMID:8264798
A:Accession: T28466
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-333 <MAS>
A:Cross-references: EMBL:L22579; NID:G623595; PIDN:AAA60776.1; PID:G438946
A:Experimental source: strain "Bangladesh-1975"
C:Function:
A:Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleoside
C:Superfamily: ribonucleoside reductase small subunit
C:Keywords: deoxyribonucleotide biosynthesis; oxidoreductase
Query Match 70.2%; Score 1279; DB 2; Length 333;
Best Local Similarity 74.7%; Pred. No. 3.5e-97;
Matches 239; Conservative 35; Mismatches 44; Indels 2; Gaps 2;
QY 32 EPLLKRSRRFVIFPIQYDINWKYKQAQSFMTAEVDLSKDLPHNKLKADKYPFISH 91
Db 16 EPLIAPNPNRFVIFPIQYDINWKYKKAASFTVVEVDISKDINDWNKLTPEKYPFIKH 75
QY 92 ILAFAASDGINVENLVERSQEVOVPEARCFYGFQILLIENVHSEMYSLLLDITVYRDPKK 151
Db 76 VLAFFAASDGINVENLAERFCTEQITEARCFYGFQMAIENHSEMYSLLLDITVYKDSNE 135
QY 152 REFLFNAIETMPYKKADWALRIADRKSTFGFGRVVAFAAAGVIFGSGFASIFWLKKR 211
Db 136 KNYLFNAIETMPCKVKADWAQKWIHD-SAGYGERLIAFAAAGVIFGSGFASIFWLKKR 194
QY 212 GLMPGLTFSNELSRDEGLHCDFACLMFQYLVNKPSEERVREIIVDAVKIEQEFLEALP 271
Db 195 GLMPGLTFSNELSRDEGLHCDFACLMFQYLVNKPSEERVREIIVDAVKIEQEFLEALP 254
QY 272 VGLIGNCMLMKQYIEFVADRLVELGFSKVOAENPFDPMENISLEGKTNFFEKRVSEY 331
Db 255 VKLIGNCENMKTYIEFVADRLVELGFSKVOAENPFDPMENISLEGKTNFFEKRVSEY 314
QY 332 QRFVMAETTDNVFTLDADF 351
Db 315 QKGMVMSQ-EDNHFSLDVDF 333
RESULT 10
H36839
C8L protein - variola virus (strain India-1967)

C:Species: variola virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: H36839
R:Blinov, V.M.
submitted to GenBank, November 1992
A:Reference number: A36859
A:Accession: H36839
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-333 <BLI>
A:Cross-references: UNIPROT:P33799; GB:X69198; NID:G456758; PIDN:CAA48969.1; PID:G297209
C:Superfamily: ribonucleoside reductase small subunit

Query Match 69.5%; Score 1266; DB 2; Length 333;
Best Local Similarity 74.1%; Pred. No. 4.1e-96;
Matches 237; Conservative 35; Mismatches 46; Indels 2; Gaps 2;

QY 32 EPLLKSRFRVIFPIQYDPDIWMYKQAQAFWTAEEVDLSKDLPHWNKLKADKPYFISH 91
DB 16 EPLAKNPNRFRVIFPIQYDIWNMYKKAASFWTVEVDISKINDWNKLTDPKYPFKH 75

QY 92 ILAPFAASDGIWNNLHVPSQEVQVPEARCFYGFQIILNHNHSEMYSLIDTYIRDPKK 151
DB 76 VLAPFAASDGIWNNLHAEFCIEVQITEARCFYGFQMAIENHSEMYSLIDTYIRKDSNE 135

QY 152 REFLNIAETMPVYKKKADWLRWADKSTGTERVAPAAVEGVFFSGSFAAIFWLKKR 211
DB 136 KNYLFNAETMPCVKKADWAKQWIDH-SAGYGERLIAFAAAGVFFSGSFAIFWLKKR 194

QY 212 GLMPGLTFSNELISRDEGLHCDPACILMFQYLVNKPSEERVREIIVDAVKIEQBFLEALP 271
DB 195 GLMPGLTFSNELISRDEGLHCDPACILMFKXLLYPPSEETVRSIITDAVSIQEFLEALP 254

QY 272 VGLIGNCILMKNQYIEFVADRLLELGVSKVFOAENPDPFMENISLEGKTNFPEKRVSY 331
DB 255 VKLIGNCEMKYTFEFVADRLLELGVSKVFOAENPDPFMENISLEGKTNFPEKRVGEY 314

QY 332 QRFVMAETTDNVFTLDADF 351
DB 315 QKMGVMSQ-EDNHSLDVDF 333

RESULT 11
T30470
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - Lymantria dispar nuclea
N:Alternate names: ribonucleotide reductase small subunit homolog
C:Species: Lymantria dispar nuclear polyhedrosis virus, IdmNPV
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T30470
R:Kuzio, J.; Pearson, M.N.; Harwood, S.H.; Funk, C.J.; Evans, J.T.; Slavicek, J.M.; Rohr
Virology 253, 17-34, 1999
A:Title: Sequence and analysis of the genome of a baculovirus pathogenic for Lymantria d
A:Reference number: Z20836; MUID:99124785; PMID:9887315
A:Accession: T30470
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-348 <XUZ>
A:Cross-references: UNIPROT:Q9YMK7; EMBL:AF081810; PIDN:AAC70306.1
C:Function:
A:Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleos
C:Superfamily: ribonucleoside reductase small subunit
C:Keywords: oxidoreductase; pyrimidine deoxynucleotide metabolism

Query Match 69.1%; Score 1258.5; DB 2; Length 348;
Best Local Similarity 70.1%; Pred. No. 1.8e-95;
Matches 246; Conservative 37; Mismatches 59; Indels 9; Gaps 3;

QY 1 MGDPERPEAGLDQDERSSSDNESEIKSNBPLKSRFRVIFPIQYDPDIWKMVYKQAQ 60
DB 7 MPAPERP-----FDPRAPAP--ERPFPDPAEPPLLRNRRFRVIFPIQYDPDMRWYMKKAE 58

QY 61 ASFWTAEEVDLSKDLPHWNKLKADKPYFISHILAPFAASDGIWNNLHVPSQEVQVPEA 120

Db 59 ASFWTVEVDLSKDTSDWERNLNDNERHFIKHVLAPFAASDGIWNNLHVPSQEVQVPEA 118
QY 121 RCFYGFQIILNHNHSEMYSLIDTYIRDPKKREFLNIAETMPVYKKKADWLRWADRK 180
Db 119 RCFYGFQIILNHNHSEMYSLIDTYIRVRSSEKRLNIAETMPCKKAEWALWIAGRE 178
QY 181 STGERVAPAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDPACILMFQ 240
Db 179 AAFGERLVAPAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDPACILFK 238
QY 241 YLVNKPSEERVREIIVDAVKIEQBFLEALPVGLIGNCILMKNQYIEFVADRLLELGVPS 300
Db 239 HLQVPSAARVREIIVDAVKIEQBFLEALPVGLIGNCILMKNQYIEFVADRLLELGVGA 298
QY 301 KVFQAEENPDPFMENISLEGKTNFPEKRVSEYQVAPVMAETTDNVFTLDADF 351
Db 299 KHYNTNPDLDFMNLISLEGKTNFPEKRVGEYQKPGVNFAXHEE-FTLDADF 348

RESULT 12

RDS28R

ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - Atlantic surf clam (fraga
N:Alternate names: ribonucleotide reductase small chain
C:Species: Spisula solidissima (Atlantic surf clam)
C:Date: 28-Dec-1987 #sequence_revision 30-Sep-1988 #text_change 09-Jul-2004
C:Accession: A22259
R:Standart, N.M.; Bray, S.J.; George, E.L.; Hunt, T.; Ruderman, J.V.
J. Cell Biol. 100, 1968-1976, 1985
A:Title: The small subunit of ribonucleotide-diphosphate reductase is encoded by one of t
A:Reference number: A22259; MUID:85207963; PMID:2987274
A:Accession: A22259
A:Molecule type: DNA
A:Residues: 1-299 <STA>
A:Cross-references: UNIPROT:P07201
C:Comment: The synthesis of this polypeptide chain is triggered by fertilization of the e
C:Function:
A:Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleos
C:Superfamily: ribonucleoside reductase small subunit
C:Keywords: deoxyribonucleotide biosynthesis; iron; metalloprotein; oxidoreductase
F;45,76,79,139,173,176/Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, His) #statu
F;83/Active site: Tyr (stable tyrosyl radical) #status predicted

Query Match 66.8%; Score 1217; DB 1; Length 299;
Best Local Similarity 77.7%; Pred. No. 3.7e-92;
Matches 233; Conservative 33; Mismatches 30; Indels 4; Gaps 3;

QY 55 MYKQAQASFMTAEVDLSKDLPHWNKLKADKPYFISHILAPFAASDGIWNNLHVPSQEQ 114

Db 1 MYKKAESFW-AEVDLSKDMAHWSLKKEKFISHVLAPFAASDGIWNNLHVPSKE 59

QY 115 VQVPEARCFYGFQIILNHNHSEMYSLIDTYIRDPKKREFLNIAETMPVYKKKADWLR 174

Db 60 VGVTEARCFYGFQIILNHNHSEMYSLIDTYIRDPKKREFLNIAETMPCKKADWLR 119

QY 175 WIADKSTGTERVAPAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDP 234

Db 120 WINDSSSYAARVAPAAVEGVFFSGSFAIFWLKKRGLMPGLTFSNELISRDEGLHCDP 179

QY 235 ACLMPQVLNKPSEERVREIIVDAVKIEQBFLEALPVGLIGNCILMKNQYIEFVADRL 294

Db 180 ACLMPQVLNKPSEERVREIIVDAVKIEQBFLEALPVGLIGNCILMKNQYIEFVADRL 239

QY 295 VELGFSKVFOAENPDPFMENISLEGKTNFPEKRVSEYQVAPVMA--ETTD-NVFTLDADF 351

Db 240 LELKCDKLYNKENPDPFMENISLEGKTNFPEKRVGEYQKMGVMSGGTGDASHFTLDADF 299

RESULT 13

T18876

hypothetical protein C03C10.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T18876

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2005, 06:36:26 ; Search time 71 Seconds
(without alignments)
2531.546 Million cell updates/sec

Title: US-10-698-228-1

Perfect score: 1821

Sequence: 1 MGDPRPEAAGLDQDERSSS.....QRFAVMAETTDNVTLADAF 351

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_Q3.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	100.0	351	Q7LG56	Q7LG56 homo sapien
2	1817	99.8	351	Q9NUW3	Q9NUW3 homo sapien
3	1725	94.7	351	Q6PEE3	Q6PEE3 mus musculus
4	1534	84.2	297	Q6A141	Q6A141 homo sapien
5	1514	83.1	299	Q75PY9	Q75PY9 homo sapien
6	1474	80.9	285	Q86VE3	Q86VE3 homo sapien
7	1454	79.8	389	R1R2_HUMAN	R1R2_HUMAN
8	1443	79.2	390	R1R2_MOUSE	R1R2_MOUSE
9	1439	79.0	386	Q6BEP1	Q6BEP1
10	1431	78.6	386	Q801Q4	Q801Q4 xenopus tro
11	1427	78.4	386	Q8AVY2	Q8AVY2 xenopus lae
12	1425.5	78.3	386	R1R2_BRARE	R1R2_BRARE
13	1421	78.0	386	R1R2_MESAU	R1R2_MESAU
14	1420.5	78.0	386	Q6D144	Q6D144
15	1399.5	76.9	378	Q6P876	Q6P876
16	1385	76.1	406	Q61P47	Q61P47 xenopus tro
17	1382.5	75.9	388	Q7PE28	Q7PE28 anophelis g
18	1382.5	75.9	426	Q7Q1F4	Q7Q1F4 anophelis g
19	1370	75.2	349	Q7ZYW0	Q7ZYW0 brachydanio
20	1358	74.6	397	Q95VP8	Q95VP8 aedes aegypt
21	1349.5	74.1	399	Q27124	Q27124 urechis cau
22	1348	74.0	393	R1R2_DROME	R1R2_DROME
23	1335	73.3	398	Q9XYN8	Q9XYN8 aedes albop
24	1331	73.1	384	R1R2_SPISO	R1R2_SPISO
25	1284	70.5	403	Q6CFU6	Q6CFU6 sparrowia li
26	1280	70.3	319	Q57175	Q57175 vaccinia vi
27	1280	70.3	319	Q76Q46	Q76Q46 variola min
28	1280	70.3	319	Q89087	Q89087 variola vir
29	1279	70.2	319	R1R2_VACCV	R1R2_VACCV
30	1279	70.2	319	Q76ZX1	Q76ZX1 vaccinia vi
31	1277	70.1	319	Q6RZQ9	Q6RZQ9 rabbitpox v

RESULT 1

ID	Q7LG56	PRELIMINARY;	PRT;	351 AA.
AC	Q7LG56	Q9NPD6; Q9NTD8;		
DT	05-JUL-2004	(TrEMBLrel. 27, Created)		
DT	05-JUL-2004	(TrEMBLrel. 27, Last sequence update)		
DT	25-OCT-2004	(TrEMBLrel. 28, Last annotation update)		
DE	Ribonucleotide reductase (Hypothetical protein DKFZp761E1312) (P53-			
DE	Inducible ribonucleotide reductase small subunit 2).			
GN	Name=p53R2; Synonyms=DKFZp761E1312;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20179179; PubMed=10716435; DOI=10.1038/35003506;			
RA	Tanaka H., Arakawa H., Yamaguchi T., Shiraishi K., Fukuda S.,			
RA	Matsui K., Takei Y., Nakamura Y.;			
RT	"A ribonucleotide reductase gene involved in a p53-dependent cell-			
RL	Nature 404:42-49 (2000).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Amgda;			
RG	The German CDNA Consortium;			
RA	Ansoerge W., Krieger S., Regiert T., Rittmueller C., Schwager B.,			
RA	Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;			
RL	Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	Ugai H., Yokoyama K.K.;			
RL	Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AB036532; BAA92493.1; -			
DR	EMBL; AB036524; BAA92493.1; JOINED.			
DR	EMBL; AB036525; BAA92493.1; JOINED.			
DR	EMBL; AB036526; BAA92493.1; JOINED.			
DR	EMBL; AB036527; BAA92493.1; JOINED.			
DR	EMBL; AB036528; BAA92493.1; JOINED.			
DR	EMBL; AB036529; BAA92493.1; JOINED.			
DR	EMBL; AB036530; BAA92493.1; JOINED.			
DR	EMBL; AB036531; BAA92493.1; JOINED.			
DR	EMBL; AB137348; CAB70703.2; -			
DR	EMBL; AB166671; BAD12267.1; -			
DR	EMBL; AB036063; BAA92434.1; -			
DR	Genew; HGNC:17296; RRM2B.			
DR	GO; GO:0004748; P:ribonucleoside-diphosphate reductase activity; IEA.			
DR	GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.			
DR	InterPro; IPR009078; Ferritin/RR like.			
DR	InterPro; IPR000358; Ribonucleotide reductase.			
DR	PFam; PF00268; Ribonucleotide reductase.			
DR	PROSITE; PS00368; RIBONUCLEOTIDE REDUCTASE.			
KW	Hypothetical protein.			
SQ	SEQUENCE 351 AA; 40736 MW; 6D008687E6F40994 CRC64;			

ALIGNMENTS

P20493 vaccinia vi
Q81lh6 ectromelia
Q89559 variola vir
Q8gm14 cowpox viru
P29883 vaccinia vi
Q8V544 monkeypox v
P87632 cowpox viru
Q8V2x8 camelpox vi
Q77525 camelpox vi
Q91fe9 vaccinia vi
P33799 variola vir
Q9C167 neurospora
Q9pxr4 variola vir
Q9ymk7 lymantria d

32 1276 70.1 319 1 R1R2_VACCC
33 1276 70.1 319 2 Q8JLH6
34 1276 70.1 319 2 Q89559
35 1274 70.0 333 2 Q8QNI4
36 1273 69.9 319 1 R1R2_VACCP
37 1271 69.8 319 2 Q8V544
38 1270 69.7 319 2 P87632
39 1268 69.6 319 2 Q8V2X8
40 1268 69.6 319 2 Q77525
41 1267 69.6 319 2 Q91FE9
42 1266 69.5 319 1 R1R2_VARV
43 1264.5 69.4 410 1 R1R2_NEUCR
44 1260 69.2 333 2 Q9PXR4
45 1258.5 69.1 348 2 Q9YMK7

Query Match 100.0%; Score 1821; DB 2; Length 351;
Best Local Similarity 100.0%; Pred. No. 6, 1e-138;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSDDTNESEIKNEEPLLRKSSRRFVFPFIQYDPDIWKMYKQAO 60
DB 1 MGDPERPEAAGLDQDERSDDTNESEIKNEEPLLRKSSRRFVFPFIQYDPDIWKMYKQAO 60
QY 61 ASFWTAEVDLSKDLPHNKLKADKEYFISHILAFPAASDGINVENLVERFSQEVQVPEA 120
DB 61 ASFWTAEVDLSKDLPHNKLKADKEYFISHILAFPAASDGINVENLVERFSQEVQVPEA 120
QY 121 RCFYGFQILLENVHSEMYSLIDTYIRDPKRELFNAIETMPYVKKADWALRWIADRK 180
DB 121 RCFYGFQILLENVHSEMYSLIDTYIRDPKRELFNAIETMPYVKKADWALRWIADRK 180
QY 181 STFGERVVAPAAVEGVFFSGSFAAIFWLRKGLMPLGLTFSNELISRDEGLHCDFACLMFQ 240
DB 181 STFGERVVAPAAVEGVFFSGSFAAIFWLRKGLMPLGLTFSNELISRDEGLHCDFACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVLVGLGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVLVGLGFS 300
QY 301 KVFOAENPDPFMEINISLEGKTNFFKRVSEYQORFAVMAETTDNVFTLDADF 351
DB 301 KVFOAENPDPFMEINISLEGKTNFFKRVSEYQORFAVMAETTDNVFTLDADF 351

RESULT 2
Q9NUN3 PRELIMINARY; PRT; 351 AA.

AC Q9NUN3
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Hypothetical protein FLJ11103.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta; DOI=10.1038/ngl1285;
RX PubMed=14702039; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Iehli S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku T., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y., Ishida S.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musahiro K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Moniyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Hishigaki H., Watanabe T., Sugiyama H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama H., Takenoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Omori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,

RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs";
RL Nat. Genet. 36:40-45(2004).
DR ENBL; AK001965; BAA92005.1; --
DR HSSP; P11157; 1XSM.
DR GO: GO:0004748; P:deoxyribonucleoside diphosphate reductase activity; IEA.
DR GO: GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonucleotidase.
DR Pfam; PF00268; Ribonucleotidase; 1.
DR PROSITE; PS00368; RIBONUC SMALL; 1.
SQ SEQUENCE 351 AA; 40704 MW; 6D009B2D59E9A323 CRC64;

Query Match 99.8%; Score 1817; DB 2; Length 351;
Best Local Similarity 99.7%; Pred. No. 1, 3e-137;
Matches 350; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSDDTNESEIKNEEPLLRKSSRRFVFPFIQYDPDIWKMYKQAO 60
DB 1 MGDPERPEAAGLDQDERSDDTNESEIKNEEPLLRKSSRRFVFPFIQYDPDIWKMYKQAO 60
QY 61 ASFWTAEVDLSKDLPHNKLKADKEYFISHILAFPAASDGINVENLVERFSQEVQVPEA 120
DB 61 ASFWTAEVDLSKDLPHNKLKADKEYFISHILAFPAASDGINVENLVERFSQEVQVPEA 120
QY 121 RCFYGFQILLENVHSEMYSLIDTYIRDPKRELFNAIETMPYVKKADWALRWIADRK 180
DB 121 RCFYGFQILLENVHSEMYSLIDTYIRDPKRELFNAIETMPYVKKADWALRWIADRK 180
QY 181 STFGERVVAPAAVEGVFFSGSFAAIFWLRKGLMPLGLTFSNELISRDEGLHCDFACLMFQ 240
DB 181 STFGERVVAPAAVEGVFFSGSFAAIFWLRKGLMPLGLTFSNELISRDEGLHCDFACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVLVGLGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVLVGLGFS 300
QY 301 KVFOAENPDPFMEINISLEGKTNFFKRVSEYQORFAVMAETTDNVFTLDADF 351
DB 301 KVFOAENPDPFMEINISLEGKTNFFKRVSEYQORFAVMAETTDNVFTLDADF 351

RESULT 3
Q6PEE3 PRELIMINARY; PRT; 351 AA.

AC Q6PEE3
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE Ribonucleotide reductase M2 B (P53 inducible).
GN Name=Rrm2b;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Srausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RT Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC058103; AAH58103.1; -
DR GO; GO:0004748; P:ribonucleoside-diphosphate reductase activity; IDA.
DR GO; GO:0009263; P:deoxyribonucleotide biosynthesis; IDA.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonuc_redctse.
DR Pfam; PF00268; Ribonuc_red_sm; 1.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
SQ SEQUENCE 351 AA; 40803 MW; 4E1259233C9C8A9 CRC64;

Query Match 94.7%; Score 1725; DB 2; Length 351;
Best Local Similarity 93.4%; Pred. No. 3.2e-130;
Matches 328; Conservative 14; Mismatches 9; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSSTTNESEIKSNEPPLKSSRRRVIPPIQVYDIDWKYKQQAQ 60
DB 1 MGDPERPEAARPEKEQGLCSSTEENVRNSNEPPLKSSRRRVIPPIQVYDIDWKYKQQAQ 60

QY 61 ASFWTAEVDLSKDLPHWNKLADEKYFISHILAFFAASDGIWNLVERFSQEVQVPEA 120
DB 61 ASFWTAEVDLSKDLPHWNKLSDEKYFISHILAFFAASDGIWNLVERFSQEVQVPEA 120

QY 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKREFLNAIETMPYVKKADWALRIADRK 180
DB 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKREFLNAIETMPYVKKADWALRIADRK 180

QY 181 STFGERVVAFVAVGVFSSGFAAFWLKRGMLPGLTFSPNLISSRDEGLHCDPACLMFQ 240
DB 181 STFGERVVAFVAVGVFSSGFAAFWLKRGMLPGLTFSPNLISSRDEGLHCDPACLMFQ 240

QY 241 YLVNKPSEERVREIIVDAVKIEQFLTEALPVGLIGMNCILMKQYIEFVADRLVVELGFS 300
DB 241 YLVNKPSEDRVREIIVDAVQIEQFLTEALPVGLIGMNCILMKQYIEFVADRLVVELGFS 300

QY 301 KVFQAEHPDFPMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 301 KIFQAEHPDFPMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 4
Q6A141 PRELIMINARY; PRT; 297 AA.
AC Q6A141
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein DKFZp686M05248.
GN Name=DKFZp686M05248;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RG The German cDNA Consortium;
RA Ottenwelder B., Obermaier B., Deutschenbauer S., Schaiipp A.,
RA Mewes H.W., Weil B., Anid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR627376; CAH10473.1; -
DR GO; GO:0004748; P:ribonucleoside-diphosphate reductase activity; IEA.
DR GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.

DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonuc_redctse.
DR Pfam; PF00268; Ribonuc_red_sm; 1.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
KW Hypothetical protein.
SQ SEQUENCE 297 AA; 34498 MW; 2B522F2CC158A02 CRC64;

Query Match 84.2%; Score 1534; DB 2; Length 297;
Best Local Similarity 99.7%; Pred. No. 5.7e-115;
Matches 296; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 55 MYKQAQASFWTAEVDLSKDLPHWNKLADEKYFISHILAFFAASDGIWNLVERFSQEV 114
DB 1 MYKQAQASFWTAEVDLSKDLPHWNKLADEKYFISHILAFFAASDGIWNLVERFSQEV 60

QY 115 VQVPEARCFYGFQILLIENVHSEMYSLIDTYIRDPKREFLNAIETMPYVKKADWALRI 174
DB 61 VQVPEARCFYGFQILLIENVHSEMYSLIDTYIRDPKREFLNAIETMPYVKKADWALRI 120

QY 175 WIADRKSTFGERVVAFVAVGVFSSGFAAFWLKRGMLPGLTFSPNLISSRDEGLHCDF 234
DB 121 WIADRKSTFGERVVAFVAVGVFSSGFAAFWLKRGMLPGLTFSPNLISSRDEGLHCDF 180

QY 235 ACLMFOYLNVNKPSEERVREIIVDAVKIEQFLTEALPVGLIGMNCILMKQYIEFVADRL 294
DB 181 ACLMFOYLNVNKPSEERVREIIVDAVKIEQFLTEALPVGLIGMNCILMKQYIEFVADRL 240

QY 295 VELGFSKVFOAENPDPFMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 241 VELGFSKVFOAENPDPFMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTLDADF 297

RESULT 5
Q75PY9 PRELIMINARY; PRT; 299 AA.
ID Q75PY9
AC Q75PY9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE P53-inducible ribonucleotide reductase small subunit 2 long form.
GN Name=P53R2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ugai H., Yokoyama K.K.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB163437; BAD11774.1; -
DR GO; GO:0004748; P:ribonucleoside-diphosphate reductase activity; IEA.
DR GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonuc_redctse.
DR Pfam; PF00268; Ribonuc_red_sm; 1.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
SQ SEQUENCE 299 AA; 34528 MW; 4705C44389EB689B CRC64;

Query Match 83.1%; Score 1514; DB 2; Length 299;
Best Local Similarity 85.2%; Pred. No. 2.3e-113;
Matches 299; Conservative 0; Mismatches 0; Indels 52; Gaps 1;

QY 1 MGDPERPEAAGLDQDERSSTTNESEIKSNEPPLKSSRRRVIPPIQVYDIDWKYKQQAQ 60
DB 1 MGDPERPEAAGLDQD----- 15

QY 61 ASFWTAEVDLSKDLPHWNKLADEKYFISHILAFFAASDGIWNLVERFSQEVQVPEA 120
DB 16 -----EVDLSKDLPHWNKLADEKYFISHILAFFAASDGIWNLVERFSQEVQVPEA 68

QY 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKREFLNAIETMPYVKKADWALRIADRK 180
DB 69 RCFYGFQILLIENVHSEMYSLIDTYIRDPKREFLNAIETMPYVKKADWALRIADRK 128


```

QY 181 STFGSRVAFVAVGVVFGSGFAAIFWLKKGKGLMPLGTFESNELISRDEGLHCDFACLMFQ 240
DB 129 STFGSRVAFVAVGVVFGSGFAAIFWLKKGKGLMPLGTFESNELISRDEGLHCDFACLMFQ 188
QY 241 YLVNKPSESRVREIIVDAVKIQEFLTEALPVLGIGMNCILMKQYIEFVADRLVVELGFS 300
DB 189 YLVNKPSESRVREIIVDAVKIQEFLTEALPVLGIGMNCILMKQYIEFVADRLVVELGFS 248
QY 301 KVFOAENPDFMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 249 KVFOAENPDFMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 299

RESULT 6
Q86YE3 PRELIMINARY; PRT; 285 AA.
AC Q86YE3
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-WAR-2004 (Tremblrel. 26, Last annotation update)
DE Similar to ribonucleotide reductase M2 polypeptide (Fragment).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL EMBL; BC042468; AAH42468.1; --
DR HSSP; P1157; LXSM.
DR GO; GO:0004748; P:Ribonucleoside-diphosphate reductase activity; IEA.
DR GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonucleotide reductase.
DR Pfam; PF00268; Ribonucleotide reductase.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
FT NON_TER 1
SQ SEQUENCE 285 AA; 33075 MW; C4E7D536479BC15B CRC64;

Query Match 80.9%; Score 1474; DB 2; Length 285;
Best Local Similarity 100.0%; Pred. No. 3.6e-110;
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 EVDLSKDLPHWNKLUKADKDYFISHILAPFAASDGIWVNLVERFSQEVQVPEARCFYGF 126
DB 1 EVDLSKDLPHWNKLUKADKDYFISHILAPFAASDGIWVNLVERFSQEVQVPEARCFYGF 60
QY 127 QILINVHSEMYSLIDTVIRDPKKREFLNFAETMPYKKADWALRWIADRKSTFGER 186
DB 61 QILINVHSEMYSLIDTVIRDPKKREFLNFAETMPYKKADWALRWIADRKSTFGER 120
QY 187 VVAFVAVGVVFGSGFAAIFWLKKGKGLMPLGTFESNELISRDEGLHCDFACLMFQYLVNKP 246
DB 121 VVAFVAVGVVFGSGFAAIFWLKKGKGLMPLGTFESNELISRDEGLHCDFACLMFQYLVNKP 180
QY 247 SEERVREIIVDAVKIQEFLTEALPVLGIGMNCILMKQYIEFVADRLVVELGFSKVFOAE 306
DB 181 SEERVREIIVDAVKIQEFLTEALPVLGIGMNCILMKQYIEFVADRLVVELGFSKVFOAE 240
QY 307 NPDFMENISLGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 241 NPDFMENISLGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 285

RESULT 7
RIR2_HUMAN
ID RIR2_HUMAN STANDARD; PRT; 389 AA.
AC P31350;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)

```

```

DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Ribonucleoside-diphosphate reductase M2 chain (EC 1.17.4.1)
DE (Ribonucleoside-diphosphate reductase small chain).
GN Name=RRM2; Synonyms=RR2;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast carcinoma;
RX MEDLINE=92329977; PubMed=1627826;
RA Pavloff N., Rivard D., Masson S., Shen S.-H., Mes-Masson A.-M.;
RT "Sequence analysis of the large and small subunits of human
RT ribonucleotide reductase.";
RL DNA Seq. 2:227-234(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21974633; PubMed=11978970;
RA Zhou B., Yen Y.;
RT "Characterization of the human ribonucleotide reductase M2 subunit
RT gene; genomic structure and promoter analyses.";
RL Cytogenet. Cell Genet. 95:52-59(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle, and Skin;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Viallon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Provides the precursors necessary for DNA synthesis.
CC -!- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate +
CC thiorodexin disulfide + H(2)O = ribonucleoside diphosphate +
CC thiorodexin.
CC -!- COFACTOR: Binds 2 iron ions per subunit (By similarity).
CC -!- PATHWAY: DNA replication pathway; first step.
CC -!- SUBUNIT: Heterodimer of a large (M1) and a small chain (M2).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- DISEASE: Ribonucleotide reductase is thought to mediate the
CC pathogenesis of the immunodeficiency of adenosine deaminase or
CC purine nucleoside phosphorylase. The deoxynucleotides that
CC accumulate in the lymphoid cells of these patients are thought to
CC feed-back inhibit ribonucleotide reductase, preventing DNA
CC replication and cell proliferation.
CC -!- MISCELLANEOUS: Two distinct regulatory sites have been defined:
CC the specificity site, which controls substrate specificity, and
CC the activity site which regulates overall catalytic activity. A
CC substrate-binding catalytic site, located on M1, is formed only in
CC the presence of the second subunit M2.
CC -!- SIMILARITY: Belongs to the ribonucleoside diphosphate reductase
CC small chain family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```


DR GO:0009262; P-deoxyribonucleotide metabolism; IDA.
 DR InterPro: IPR000078; Ferritin/RR like.
 DR InterPro: IPR000358; Ribonuc redctase.
 DR Pfam: PF00268; Ribonuc_red_sm; 1.
 DR PROSITE: PS00368; RIBORED_SMALL; 1.
 KW 3D-structure; DNA replication; Iron; Metal-binding; Oxidoreductase.
 DR METAL 139 139 Iron 1.
 FT METAL 170 170 Iron 1 and 2.
 FT METAL 173 173 Iron 1.
 FT METAL 233 233 Iron 2.
 FT METAL 267 267 Iron 1 and 2.
 FT METAL 270 270 Iron 2.
 FT ACT_SITE 177 177 By similarity.
 FT HELIX 68 70
 FT TURN 72 74
 FT HELIX 89 99
 FT TURN 100 101
 FT HELIX 105 107
 FT TURN 112 113
 FT HELIX 114 118
 FT TURN 119 119
 FT HELIX 122 148
 FT TURN 149 149
 FT HELIX 150 153
 FT HELIX 157 184
 FT HELIX 188 195
 FT TURN 196 196
 FT HELIX 197 200
 FT HELIX 202 204
 FT HELIX 205 216
 FT TURN 218 219
 FT HELIX 222 235
 FT TURN 236 237
 FT HELIX 238 249
 FT TURN 250 251
 FT HELIX 254 279
 FT TURN 280 280
 FT HELIX 287 306
 FT TURN 307 308
 FT HELIX 311 314
 FT TURN 315 315
 FT HELIX 318 335
 FT TURN 336 337
 FT HELIX 349 351
 SQ SEQUENCE 390 AA; 45095 MW; AC7ACC4PAP8A4A2F CRC64;
 Query Match 79.2%; Score 1443; DB 1; Length 390;
 Best Local Similarity 78.9%; Pred. No. 1.6e-107;
 Matches 269; Conservative 40; Mismatches 26; Indels 6; Gaps 1;
 QY 17 RSSDTSNESEIKS-----NEEPLLRKSSRRFVIFPIQYDPIWKMVKQAASFTWTAEEVD 70
 DB 50 RIFQDSAELESKAPTNPVSVEDELLRENPRFRVFPPIEYHDIWQMYKKAASEFTWTAEEVD 109
 QY 71 LSKDLPNNKLADEKYFISHILAPFAASDGIWNLVERFSEVQVPEARCFYGFQILI 130
 DB 110 LSKDLPNNKLADEKYFISHILAPFAASDGIWNLVERFSEVQVPEARCFYGFQI 169
 QY 131 ENVHSEMYSLLDITYIRDPKREFFNAIETMPYVKKADWALRWADRKSTFGERVVAF 190
 DB 170 ENHSEMYSLLDITYIKDPKREFFNAIETMPYVKKADWALRWADRKSTFGERVVAF 229
 QY 191 AAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDPACLMFQYLVNKPSE 250
 DB 230 AAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDPACLMFQYLVNKPSE 289
 QY 251 VREIIVDAVKIQEFLTEALPVGLIGMNCILMKQYIEFVADRLLLVELGFSKVFQANPFD 310
 DB 290 VREIITNAVKIQEFLTEALPVGLIGMNCILMKQYIEFVADRLLLVELGFSKVFQANPFD 349
 QY 311 FMENISLEKTNPFKRVSEYQRFVAAETITDNVFTLDADP 351
 DB 350 FMENISLEKTNPFKRVSEYQRFVAAETITDNVFTLDADP 390

RESULT 9

Q68EP1 PRELIMINARY; PRT; 386 AA.
 ID Q68EP1
 AC Q68EP1
 DT 25-OCT-2004 (TREMBlrel. 28, Created)
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
 DE Hypothetical protein.
 OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=8364;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Matek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX Klein S., Gerhard D.S.;
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC080161; AAH80161.1; -
 DR InterPro: IPR009078; Ferritin/RR like.
 DR InterPro: IPR000358; Ribonuc redctase.
 DR Pfam: PF00268; Ribonuc_red_sm; 1.
 DR PROSITE: PS00368; RIBORED_SMALL; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 386 AA; 44540 MW; 277DF859B6053824 CRC64;
 Query Match 79.0%; Score 1439; DB 2; Length 386;
 Best Local Similarity 78.1%; Pred. No. 3.4e-107;
 Matches 268; Conservative 39; Mismatches 36; Indels 0; Gaps 0;
 QY 9 AAGLDQDERSSDTSNESEIKSNEEPLLRKSSRRFVIFPIQYDPIWKMVKQAASFTWTAEE 68
 DB 44 ARNIFQEAETTKSKAPDKPRIQDEPFLKNDPNRFRVFPPIQYHDIWQMYKKAASEFTWTAEE 103
 QY 69 VDLKSLPHNNKLADEKYFISHILAPFAASDGIWNLVERFSEVQVPEARCFYGFQI 128
 DB 104 VDLKSLPHNNKLADEKYFISHILAPFAASDGIWNLVERFSEVQVPEARCFYGFQI 163
 QY 129 LIENVHSEMYSLLDITYIRDPKREFFNAIETMPYVKKADWALRWADRKSTFGERVV 188
 DB 164 AMENIHSEMYSLLDITYIKDPKREFFNAIETMPYVKKADWALRWADRKSTFGERVV 223
 QY 189 AFAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDPACLMFQYLVNKPSE 248
 DB 224 AFAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDPACLMFQYLVNKPSE 283
 QY 249 ERVREIIVDAVKIQEFLTEALPVGLIGMNCILMKQYIEFVADRLLLVELGFSKVFQANP 308

```

Db 284 ERVQLITDAVQEQEFLTEALPVNLGMNCTLMQVIEFVADRLLELGLFNKVFRTATNP 343
Qy 309 FDFMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 351
Db 344 FDFMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 386

RESULT 10
Q80104
ID Q80104 PRELIMINARY; PRT; 386 AA.
AC Q80104;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Rm2-prov protein
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC047975; AAH47975.1; -.
DR HSSP; P11157; 1XSM.
DR GO; GO:0004748; F:Ribonucleoside-diphosphate reductase activity; IEA.
DR GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonucleotide.
DR Pfam; PF00268; Ribonuc red sm; 1.
DR PROSITE; PS00368; RIBONUC_RED_SML; 1.
DR PROSITE; PS00368; RIBONUC_RED_SML; 1.
SQ SEQUENCE 386 AA; 44595 MW; 3C12900A1CA7D54A CRC64;

Query Match 78.6%; Score 1431; DB 2; Length 386;
Best Local Similarity 79.5%; Pred. No. 1.5e-106;
Matches 267; Conservative 38; Mismatches 27; Indels 4; Gaps 1;

```

```
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041209; AAH41209.1; -.
DR HSP; P11157; IXSM.
DR GO; GO:0004748; F:Ribonucleoside-diphosphate reductase activity; IEA.
DR GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; RibonucL reductase.
DR Pfam; PF00268; Ribonuc_red_sm; 1.
DR Pfam; PF00268; Ribonuc_red_sm; 1.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
DR SEQUENCE 386 AA; 44635 MW; 4635FCC9BB18EA32 CRC64;

Query Match 78.4%; Score 1427; DB 2; Length 386;
Best Local Similarity 79.2%; Pred. No. 3.1e-106;
Matches 266; Conservative 36; Mismatches 30; Indels 4; Gaps 1;

QY 20 SDTNSSEIKSN-----EPELLKSSRRFVPIPIQYDPDIWKMYKQAQSFMTAEVDLSKOL 75
DB 51 TETPKSKAPKNPRLEDEPLKDKNPHRFVPIPIQYHDIWQYKKAESFMTAEVDLSKOL 110
QY 76 PHWNKLKADKPKYFISHILAFFAASDGIWNLVERFSQVQVPEARCFYGFQILLENVHS 135
DB 111 QHWESLKKEEKYFISHVLFAFPAASDGIWNLVERFSQVQVTEARCFYGFQIAMENHS 170
QY 136 EMYSLLIDTYIRDPKREPLFNAIETMPYVKKADWALRWIADRKSTFGERVVAFPAVEG 195
DB 171 EMYSLLIDTYIRDPKREPLFNAIETPCVKKADWALRWIADRKSTFGERVVAFPAVEG 230
QY 196 VFFSGSFAAIFWLKRGKGLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEERVREII 255
DB 231 IFFSGSFAAIFWLKRGKGLMPGLTFSNELISRDEGLHCDFACLMFKLRKPSSEERVVELI 290
QY 256 VDAVKIEQBFLEALPVGLIGMNCILMKQYIEFVADRLLELVELGFSKVFOENPFPFMEI 315
DB 291 TDAVQIEQBFLEALPVNIGMNCILMKQYIEFVADRLLELVELGFSKVFOENPFPFMEI 350
QY 316 SLEGTNFFKEKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 351 SLEGTNFFKEKRVSEYQRFVAVMAETTDNVFTLDADF 386

RESULT 12
RIR2 BRARE
ID RIR2 BRARE STANDARD; PRT; 386 AA.
AC P19733;
DT 15-DEC-1998 (Rel. 37, Created)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Ribonucleoside-diphosphate reductase M2 chain (EC 1.17.4.1)
DE (Ribonucleotide reductase protein R2 class I).
GN Name=rnm2; Synonyms=r2;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OC NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97137859; PubMed=8983196;
RA Mathews C.Z., Sjoeborg B.-M., Karlsson M.;
RT "Cloning and sequencing of cDNAs encoding ribonucleotide reductase
RT from zebrafish Danio rerio."
RL Mol. Mar. Biol. Biotechnol. 5:284-287(1996).
CC -!- FUNCTION: Provides the precursors necessary for DNA synthesis.
CC -!- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate +
CC thioedoxin disulfide + H(2)O = ribonucleoside diphosphate +
CC thioedoxin.
CC -!- COFACTOR: Binds 2 iron ions per subunit (By similarity).
CC -!- PATHWAY: DNA replication pathway; first step.
CC -!- SUBUNIT: Heterodimer of a large (M1) and a small chain (M2).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the ribonucleoside diphosphate reductase
```

359 KTNFFEKRVGEYQRMGVMS-----NSFTLDADF 386

RESULT 14

Q6DI44 PRELIMINARY; PRT; 386 AA.

ID Q6DI44

AC Q6DI44;

DT 25-OCT-2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Ribonucleotide reductase M2 polypeptide.

GN Names:rm2;

OS Brachydanio rerio (zebrafish) (Danio rerio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;

OC Cyprinidae; Danio.

OX NCBI_TaxID=7955;

RN [1]

RN SEQUENCE FROM N.A.

RP RP

RC TTSUVE=Embryo;

RC MEDLINE=223388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heih F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Udutin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Pahay J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywicki M.I., Skalka U., Smulius D.E., Schnrner A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";

RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RL [2]

RL SEQUENCE FROM N.A.

RP RP

RC TTSUVE=Embryo;

RC Strausberg R.;

RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC075746; AAH75746.1; -

DR GO; GO:0004748; F:ribonucleoside-diphosphate reductase activity; IEA.

DR GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.

DR InterPro; IPRO09078; Ferritin/RR like.

DR InterPro; IPRO00358; Ribonuclease.

DR Pfam; PF00268; Ribonuc_red_sm; 1.

DR PROSITE; PS00368; RIBOED_SMALL; 1.

SQ SEQUENCE 386 AA; 44567 MW; 2CCACD4424DD1566 CRC64;

Query Match 78.0%; Score 1420.5; DB 2; Length 386;

Best Local Similarity 78.5%; Pred. No. 1e-105;

Matches 267; Conservative 37; Mismatches 31; Indels 5; Gaps 1;

QY 17 RSSSDTNSSEIKS-----NEEPLRKSSRRFVFPIQYDPDIWKMYKQAQASFWTAEEVDL 71

DB 47 RKTFDESEGSKAKKGAVEEPELLKNPHRFVFPIQYHDIWQYKKAASFWTAEEVDL 106

QY 72 SKDLPTWVKLKADEKVFISHILAFPAASDGIWNVLENVRFSEQVDFEARCFYGFQILLIE 131

DB 107 SKDLQHWDSLKDREYFISHVLFAAASDGIWNVLENVRFTEQVDFEARCFYGFQIAE 166

QY 132 NVHSEMYSLIIDTYIRDPKKREPLFNAITMPVYKKADWALRWADRKSTFGERVVAF 191

DB 167 NVHSEMYSLIIDTYIRKSKREPLFNAITMPVCKKADWALRWIGDKNARYGERVVAF 226

QY 192 AVEGVFFSGSFAAIFWLKXKRGMLPGITFSNLIISRDEGLHCDFACLMFQYLVNKPSEEV 251

Db 227 AVEGIFFGSPASIFWLKKRGLMPLGTFNSNELISRDEGLHCDPACLMFKHLISKPSSETV 286
QY 252 RELIIVDAVKIEQBELTEALPVGLIGMNCILMKQYIEFVADRLVGLGFSKVOAENPDF 311
Db 287 KKLIMNAVEIEQBELDALPVKLGIMNCILMKQYIEFVADRLVGLGFSKVOAENPDF 346
QY 312 MENISLEGKTNPFKEKRVSEYQRFVAVMAETDNDVFTLDADF 351
Db 347 MENISLEGKTNPFKEKRVSEYQRFVAVMAETDNDVFTLDADF 386

RESULT 15
Q6P876
ID Q6P876 PRELIMINARY; PRT; 378 AA.
AC Q6P876;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE Hypothetical protein MGC75900.
GN Name=MGC75900;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8364;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haiech F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalusz D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Gerhard D.S.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC061353; AAH61353.1; -
DR GO; GO:0004748; F:ribonucleoside-diphosphate reductase activity; IEA.
DR GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; RibonucL_redctse.
DR Pfam; PF00268; Ribonuc red sm; 1.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
KW Hypothetical protein.
SQ SEQUENCE 378 AA; 43939 MW; 538D477FBB7CDB65 CRC64;

Search completed: October 30, 2005, 06:57:37
Job time : 72 secs

Query Match 76.9%; Score 1399.5; DB 2; Length 378;
Best Local Similarity 76.6%; Pred. No. 4.9e-104;
Matches 262; Conservative 45; Mismatches 28; Indels 7; Gaps 1;
QY 17 RSSSDNESEI-----KSNEPLRKSSRRFVIPPQYDIWKYQAQASFRTAEV 69
Db 37 RTSRGTHKQWQCSLKDPRIQDEPLLRDNPGRFVILPIEYVHDIWQYKAEASFRTAEV 96
QY 70 DLKSDLPHNKLKADKYPFISHILAFPAASDGIWNLVERFSQEVQVPCRCFYGQIL 129

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2005, 06:50:41 ; Search time 25 Seconds
(without alignments)
1048.073 Million cell updates/sec

Title: US-10-698-228-1
Perfect score: 1821
Sequence: 1 MGDPERPEAGLDQDERSSS.....QRFAVMAETDNTVFTLDADF 351

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5A-COMB.pep: *
2: /cgn2_6/ptodata/1/iaa/5B-COMB.pep: *
3: /cgn2_6/ptodata/1/iaa/6A-COMB.pep: *
4: /cgn2_6/ptodata/1/iaa/6B-COMB.pep: *
5: /cgn2_6/ptodata/1/iaa/PCTUS-COMB.pep: *
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1821	100.0	351	4	US-10-019-733-1
2	1454	79.8	389	4	US-10-019-733-5
3	1454	79.8	389	4	US-09-949-016-6016
4	1454	79.8	453	4	US-09-949-016-7896
5	1131	62.1	411	4	US-09-248-796A-18017
6	1072.5	58.9	274	1	US-08-307-499-19
7	1072.5	58.9	274	3	US-09-299-268-19
8	974.5	53.5	322	1	US-08-136-743B-2
9	969.5	53.2	360	4	US-09-248-796A-18018
10	768	42.2	147	3	US-08-908-223-358
11	599	32.9	149	4	US-09-270-767-46236
12	495	27.2	323	4	US-09-902-540-10006
13	381.5	21.0	445	4	US-09-328-352-4714
14	377.5	20.7	476	4	US-09-252-991A-19358
15	335	18.4	365	4	US-09-198-452A-1059
16	335	18.4	380	4	US-09-438-185A-986
17	298	16.4	314	3	US-09-080-044-3
18	298	16.4	314	3	US-09-531-857A-3
19	294	16.1	56	3	US-08-905-223-362
20	220	12.1	382	4	US-09-543-681A-4239
21	203.5	11.2	412	4	US-09-488-039A-8180
22	201.5	11.1	401	4	US-09-540-236-3566
23	200	11.0	325	4	US-09-107-532A-4443
24	195.5	10.7	320	3	US-09-092-437-2
25	195.5	10.7	320	4	US-09-583-110-3858
26	195.5	10.7	320	4	US-09-107-433-3136
27	195	10.7	339	4	US-09-543-681A-5069

28	169.5	9.3	323	4	US-09-042-775-2	Sequence 2, Appli
29	167	9.2	343	3	US-09-134-001C-5484	Sequence 5484, Ap
30	163	9.0	334	4	US-09-134-000C-5726	Sequence 5726, Ap
31	161	8.8	311	4	US-09-710-279-1260	Sequence 1260, Ap
32	157	8.6	324	2	US-08-813-940-2	Sequence 2, Appli
33	156	8.6	391	4	US-09-489-039A-10121	Sequence 10121, A
34	153	8.4	104	4	US-09-230-485-5	Sequence 5, Appli
35	138	7.6	321	2	US-08-813-940-6	Sequence 6, Appli
36	127	7.0	52	3	US-08-917-299-21	Sequence 21, Appl
37	127	7.0	52	3	US-09-422-662-21	Sequence 21, Appl
38	127	7.0	52	4	US-09-730-763-21	Sequence 21, Appl
39	127	7.0	52	4	US-09-429-370-21	Sequence 21, Appl
40	103	5.7	555	3	US-09-173-151A-32	Sequence 32, Appl
41	100.5	5.5	382	4	US-09-134-000C-5889	Sequence 5889, Ap
42	99	5.4	233	3	US-08-933-825-2	Sequence 2, Appli
43	98.5	5.4	896	4	US-09-270-767-46130	Sequence 46130, A
44	95	5.2	233	3	US-09-161-662-2	Sequence 2, Appli
45	93.5	5.1	3830	4	US-09-693-205A-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-10-019-733-1
; Sequence 1, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619MOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 1
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-1

Query Match 100.0%; Score 1821; DB 4; Length 351;
Best Local Similarity 100.0%; Pred. NO. 4e-189;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MGDPERPEAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPQYDPDWKMYKQAO	60
DB	1	MGDPERPEAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPQYDPDWKMYKQAO	60
QY	61	ASFTAAEEVDLSKDLPHWNKLKADKEYFISHILAFFAASDGI VNNELVERFSQEVQVPEA	120
DB	61	ASFTAAEEVDLSKDLPHWNKLKADKEYFISHILAFFAASDGI VNNELVERFSQEVQVPEA	120
QY	121	RCFYGQIILNIENHSEMYSLIDITYINDPKKREFLNAIETMPYVKKADWALRWIADRK	180
DB	121	RCFYGQIILNIENHSEMYSLIDITYINDPKKREFLNAIETMPYVKKADWALRWIADRK	180
QY	181	STFGERVVAFAAVEGVFFSGSFAAI PWLKKRGLMPGLTFSENELISRDEGLHCDPACLMFQ	240
DB	181	STFGERVVAFAAVEGVFFSGSFAAI PWLKKRGLMPGLTFSENELISRDEGLHCDPACLMFQ	240
QY	241	YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGNCILMKQYIEFVADRLVLVGLGFS	300
DB	241	YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGNCILMKQYIEFVADRLVLVGLGFS	300
QY	301	KVFAENPDFMFENISLEGKTNFFEKRVSEYQRPANVAETDNTDVTLDADF	351

Db 301 KVFOAENPPDFMENISLEGKTNFFKRVSEYQRFVMAETTDNVFTLDADF 351

RESULT 2

US-10-019-733-5
; Sequence 5, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 5
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-5

Query Match 79.8%; Score 1454; DB 4; Length 389;
Best Local Similarity 83.8%; Pred. No. 3.3e-149;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY	31	EEPLLKSSRRFVIFPIQYPDIMWKYKQAASFTAEVDLSKDLPHWNKLKADEKYFIS	90
Db	69	DEPLLRENPRFVIFPIEYHDIWMYKKAASFTAEVDLSKDIQHWESLKPEERYFIS	128
QY	91	HILAFFAASDGIYVNVNLFVSQVQVPEARCFYGFQIILNIHVHSEMYSLIIDTYIRDPK	150
Db	129	HVLAFFAASDGIYVNVNLFVSQVQVTEARCFYGFQIAMENIHSEMYSLIIDTYIKDPK	188
QY	151	KREFLFNAIETMPYVKKADWALRIADRKSTFGERVVAFVAAVEGFFSGSFAAIFWLKK	210
Db	189	EREFLFNAIETMPCVKKADWALRWIGDKATYGERVVAFAAVEGFFSGSPASIFWLKK	248
QY	211	RGMLPGLTFSNELISRDEGLHCDPACLMFQYLVNKPSEERVREIIVDAVKIEQEFLEAL	270
Db	249	RGMLPGLTFSNELISRDEGLHCDPACLMFKHLVHKPSEERVREIINAVRIEQEFLEAL	308
QY	271	PVGLIGMNCILMKQYIEFVADRLVLGFSKVFOAENPPDFMENISLEGKTNFFKRVSE	330
Db	309	PVKLIGMNCILMKQYIEFVADRLMLGLGFSKVFRVENPPDFMENISLEGKTNFFKRVGE	368
QY	331	YQRFVMAETTDNVFTLDADF 351	
Db	369	YQRMGMSSPTENSFTLDADF 389	

RESULT 3

US-09-949-016-6016
; Sequence 6016, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03

Query Match 79.8%; Score 1454; DB 4; Length 453;
Best Local Similarity 83.8%; Pred. No. 4.2e-149;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY	31	EEPLLKSSRRFVIFPIQYPDIMWKYKQAASFTAEVDLSKDLPHWNKLKADEKYFIS	90
Db	133	DEPLLRENPRFVIFPIEYHDIWMYKKAASFTAEVDLSKDIQHWESLKPEERYFIS	192
QY	91	HILAFFAASDGIYVNVNLFVSQVQVPEARCFYGFQIILNIHVHSEMYSLIIDTYIRDPK	150

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6016
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6016

QY	31	EEPLLKSSRRFVIFPIQYPDIMWKYKQAASFTAEVDLSKDLPHWNKLKADEKYFIS	90
Db	69	DEPLLRENPRFVIFPIEYHDIWMYKKAASFTAEVDLSKDIQHWESLKPEERYFIS	128
QY	91	HILAFFAASDGIYVNVNLFVSQVQVPEARCFYGFQIILNIHVHSEMYSLIIDTYIRDPK	150
Db	129	HVLAFFAASDGIYVNVNLFVSQVQVTEARCFYGFQIAMENIHSEMYSLIIDTYIKDPK	188
QY	151	KREFLFNAIETMPYVKKADWALRIADRKSTFGERVVAFVAAVEGFFSGSFAAIFWLKK	210
Db	189	EREFLFNAIETMPCVKKADWALRWIGDKATYGERVVAFAAVEGFFSGSPASIFWLKK	248
QY	211	RGMLPGLTFSNELISRDEGLHCDPACLMFQYLVNKPSEERVREIIVDAVKIEQEFLEAL	270
Db	249	RGMLPGLTFSNELISRDEGLHCDPACLMFKHLVHKPSEERVREIINAVRIEQEFLEAL	308
QY	271	PVGLIGMNCILMKQYIEFVADRLVLGFSKVFOAENPPDFMENISLEGKTNFFKRVSE	330
Db	309	PVKLIGMNCILMKQYIEFVADRLMLGLGFSKVFRVENPPDFMENISLEGKTNFFKRVGE	368
QY	331	YQRFVMAETTDNVFTLDADF 351	
Db	369	YQRMGMSSPTENSFTLDADF 389	

RESULT 4

US-09-949-016-7896
; Sequence 7896, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7896
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7896

Query Match 79.8%; Score 1454; DB 4; Length 453;
Best Local Similarity 83.8%; Pred. No. 4.2e-149;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

Db 193 HVLAFPAASDGVNENLVRFSEVQI TEARCFYQIAMENIHSEMYSLIIDTYIKDPK 252
Qy 151 KREELFNAETMPYKKADWALWTRADSKSPGERVAFVAAVEGFFSGSPAIFWLKK 210
Db 253 REELFNAETMPYKKADWALWTRADSKSPGERVAFVAAVEGFFSGSPAIFWLKK 312
Qy 211 RGLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEBVEREIIIVDAVKIEQEFLEAL 270
Db 313 RGLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEBVEREIIIVDAVKIEQEFLEAL 372
Qy 271 PVGLIGMNCILMKQIEFVADRLVELGFSKVQFAENPFDPMENISLEGKTNFFEKRVSE 330
Db 373 PVKLGIMNCILMKQIEFVADRLVELGFSKVQFAENPFDPMENISLEGKTNFFEKRVSE 432
Qy 331 YORFAVMAETTDNVTLDAF 351
Db 433 YORGMVMSPTENSFTLDAF 453

RESULT 5
US-09-248-796A-18017
; Sequence 18017, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18017
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18017

Query Match 62.1%; Score 1131; DB 4; Length 411;
Best Local Similarity 62.7%; Pred. No. 4.4e-114;
Matches 210; Conservative 52; Mismatches 67; Indels 6; Gaps 3;
Qy 23 NESEIKSNE--EPLLRKSSRRFVPIQYDPDIWKMYKQAQASFTWAEVDLSKDLPHW-N 79
Db 77 NRHLLKQESNEPLITPDKTRHTIPIKYPELWQFYKSLASFVTAEBLDSLKDLDDWN 136
Qy 80 KKADEKYPISHILAFPAASDGVNENLVRFSEVQVPEARCFYQIILNENHSEMY 139
Db 137 KMNENERFISRLVAFPAASDGVNENLVNFCAEVQIPEAKLVYKFQIMNENHSETYS 196
Qy 140 LIIDTYIDPKKRELEFNAETMPYKKADWALWTRADSKSPGERVAFVAAVEGFFS 199
Db 197 LIETIYFKDPEADFLNADINPIPKKADWALWTRADSKSPGERVAFVAAVEGFFS 256
Qy 200 GSFAAIFWKKRGLMPLTFSNELISRDEGLHCDFACLMFQYLVNKPSEBVEREIIIVDAV 259
Db 257 GSFAAIFWKKRGLMPLTFSNELISRDEGLHCDFACLMFQYLVNKPSEBVEREIIIVDAV 316
Qy 260 KIEQEFLEALPVGLIGMNCILMKQIEFVADRLVELGFSKVQFAENPFDPMENISLEG 319
Db 317 DIEKEYFDALPVSLGIMNCILMKQIEFVADRLVELGFSKVQFAENPFDPMENISLEG 376
Qy 320 KTNPEKRVSEYORFAVMAET---TDNVTLDAF 351
Db 377 KTNPEKRVSDYQKAGVMEKVENKNEKTGLPQDF 411

RESULT 6
US-08-307-499-19

; Sequence 19, Application US/08307499
; Patent No. 5651972
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,499
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF35.1.FWCCI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 274 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-307-499-19

Query Match 58.9%; Score 1072.5; DB 1; Length 274;
Best Local Similarity 75.6%; Pred. No. 5.4e-108;
Matches 204; Conservative 27; Mismatches 38; Indels 1; Gaps 1;
Qy 32 EPLLKSSRRFVPIQYDPDIWKMYKQAQASFTWAEVDLSKDLPHWKNLKADEKYPISH 91
Db 2 EPILOESDSRFVPIQYHDWIKMYKQSVASFVVEVDLSKDLDDWDLTKDEKYPISH 61
Qy 92 ILAFAAASDGVNENLVRFSEVQVPEARCFYQIILNENHSEMYSLIIDTYIDPK 151
Db 62 ILAFAAASDGVNENLVRFSEVQVPEARCFYQIILNENHSEMYSLIIDTYIDPK 121
Qy 152 REPLFNAETMPYKKADWALWTRADSKSPGERVAFVAAVEGFFSGSPAIFWLKKR 211
Db 122 KHLFNAETMECVKKKADWARKWISNK-VYGERVAFVAAVEGFFSGSPAIFWLKKR 180
Qy 212 GLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEBVEREIIIVDAVKIEQEFLEALP 271
Db 181 GLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEBVEREIIIVDAVKIEQEFLEALP 240
Qy 272 VGLIGMNCILMKQIEFVADRLVELGFSKV 301

Db 241 VDLIGNCCLMSQYIEFVADRLLTGCEK 270

RESULT 7

US-09-299-268-19

; Sequence 19, Application US/09299268

; Patent No. 6217882

; GENERAL INFORMATION:

; APPLICANT: Moyer, Richard W.

; APPLICANT: Vi uella, Eladio

; APPLICANT: Gibbs, E.P.J.

; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a

; TITLE OF INVENTION: Live Vaccine Vector

; NUMBER OF SEQUENCES: 60

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: David R. Saliwanchik

; STREET: 2421 N.W. 41st Street, Suite A-1

; CITY: Gainesville

; STATE: Florida

; COUNTRY: U.S.A.

; ZIP: 32606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/299,268

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/901,127

; FILING DATE:

; APPLICATION NUMBER: US 07/908,241

; FILING DATE: 1-JUL-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/908,630

; FILING DATE: 29-JUN-1992

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/342,212

; FILING DATE: 21-APR-1992

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Saliwanchik, David R.

; REGISTRATION NUMBER: 31,794

; REFERENCE/DOCKET NUMBER: UF35.1.FWCC1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 904-375-8100

; TELEFAX: 904-372-5800

; INFORMATION FOR SEQ ID NO: 19:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 274 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-09-299-268-19

Query Match 58.9%; Score 1072.5; DB 3; Length 274;

Best Local Similarity 75.6%; Pred. No. 5.4e-108;

Matches 204; Conservative 27; Mismatches 38; Indels 1; Gaps 1;

QY 32 EPLKSSRRFVFPFIQYDDIWKYKQASFTWAEVDLSKDLPHNKLKADEKFIH 91

Db 2 EPILOESDRFVFPFIKTHDIWKYKQSVASFTWVEVDLSKDLDDWDKLTDEKFIKH 61

QY 92 ILAFAASDGIWENLVERFSQVPEARCFYGFQILLIENHSEMYSLIIDTYIRDPKK 151

Db 62 ILAFASSDGIWENLAEAFYVDQSEARCFYGFQIAWENHSEMYSLIIDTYVRDNIE 121

QY 152 REFLFNAITETMPYVKKADWALRWIADRKSTFGERVVAFAAVEGVFSGSFAAIFWLK 211

Db 122 KHLFNAITETMECVKKADWARKWISSNK-VYGERVVAFAAVEGIFSGSFAAIFWIKR 180

QY 212 GLMPGLTFSNELISRDEGLHCDFACLMFOYLVNKPSEBVRREIIVDAVKIQEFLTEALP 271

Db 181 GLMPGLTFSNELISRDEGLHCDFACLMFKHLLHPPSPKEVITSIIIDAVNIKEFLTVAIP 240

QY 272 VGLIGNCILMKQYIEFVADRLLTGCEK 301

Db 241 VDLIGNCCLMSQYIEFVADRLLTGCEK 270

RESULT 8

US-08-136-743B-2

; Sequence 2, Application US/08136743B

; Patent No. 5459063

; GENERAL INFORMATION:

; APPLICANT: Barry S. Cooperman, Harvey Rubin,

; APPLICANT: Jerome Salem, and Alison L. Fisher

; TITLE OF INVENTION: "Plasmodium falciparum Ribonu-

; TITLE OF INVENTION: cleotide Reductase, DNA Sequences Therefor and Peptide inhibit

; TITLE OF INVENTION: Thereof"

; NUMBER OF SEQUENCES: 67

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: The University of Pennsylvania

; STREET: Suite 330

; STREET: 3700 Market Street

; CITY: Philadelphia

; STATE: Pennsylvania

; COUNTRY: U.S.A.

; ZIP: 19104-3246

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: WordPerfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/136,743B

; FILING DATE: 10/14/93

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Monaco, Daniel A.

; REGISTRATION NUMBER: 30,480

; REFERENCE/DOCKET NUMBER: 3957-10

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (215) 568-8383

; TELEFAX: (215) 568-5549

; TELEX: NO. 5459063e

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 322 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

US-08-136-743B-2

Query Match 53.5%; Score 974.5; DB 1; Length 322;

Best Local Similarity 59.7%; Pred. No. 3e-97;

Matches 191; Conservative 51; Mismatches 75; Indels 3; Gaps 2;

QY 34 LIRKSSRRFVFPFIQYDDIWKYKQASFTWAEVDLSKDLPHNKLKADEKFIH 93

Db 4 ILNKSDRFTLYPILYDPVFPFYKKAESFTWAEIDYSSDLKDFEKLNENEKFIKHVL 63

QY 94 AFAASDGIWENLVERFSQVPEARCFYGFQILLIENHSEMYSLIIDTYIRDPKRE 153

Db 64 AFAASDGIWENLAVSFLEVOITAKFYFQIAVENHSEMYSLIIDTYIRDPKRE 123

QY 154 FLFNAITETMPYVKKADWALRWIADRKSTFGERVVAFAAVEGVFSGSFAAIFWLK 213

Db 124 NLFALENTPAVKALWAAKWIINDTNS-FAERIVANACVEGIFSGSFAAIFWK 182

QY 214 MPGLTFSNELISRDEGLHCDFACLMFOYLVNKPSEBVRREIIVDA--VKIQEFLTEALP 271

Db 183 LHGLTFSNELISRDEGLHCDFACLMFOYLVNKPSEBVRREIIVDA--VKIQEFLTEALP 242

```
QY 272 VGLIGNCILMKQYIEFVADRLLVELGFSKVFOAENPDPFPMENISLEGKTNPFKEKRVSEY 331
Db 243 CDLIGNSLMSQTIIEFVADRLLLECLGCKSIKPHSKNPFNMWDKISLQKTNPFKEKRVADY 302
QY 332 QRPVMAETTDNVFTLDADP 351
Db 303 QKSGVNAQRKDHVFLCLNTEF 322

RESULT 9
US-09-248-796A-18018
; Sequence 18018, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18018
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18018

Query Match 53.2%; Score 969.5; DB 4; Length 360;
Best Local Similarity 62.2%; Pred. No. 1.2e-96;
Matches 179; Conservative 53; Mismatches 53; Indels 3; Gaps 2;

QY 5 ERPEAAGLDQDERSSDTNESEIK--SNBEPILKSSRRFVIPPPIQYPDIMWKYKQAQAS 62
Db 72 EKDDSLKXHQDFLAKHVKHRLKQLAEPLLVENKRRYVMPPIRYHINWPKYKAEAS 131
QY 63 FWTAEVDLSKDLPHW-NKLKADEKYFISHILAFFAASDGI VNNENLVRFSEVQVPEAR 121
Db 132 FWTAEIDLKDLDDNNKLNENERYFISKVLAFFAASDGI VNGENLNFSTEVQLPEAK 191
QY 122 CYGFGQILLENVHSEMYSLIIDTYIRDPKREPLFNAIETMPYVKKKADWALRWIADRS 181
Db 192 SYGFGQIMNTHSETYSLLIETIKDPQEAADYLFNAIANIPCIQKKADWAIKWIQDDEA 251
QY 182 TFGERVAFPAVEGVFFSGSFAAIFWLKKGRLMPLGLTFSGNELISRDGLHCDPACILMFOY 241
Db 252 LYGERLVAFPAVEGVFFSGSFAIFWLKKGRLMPLGLTFSGNELICRDEGLHTDPACLLFH 311
QY 242 LVNKPSEERREIIVDAVKIEQFLEALPVGILGNMNCILMKQYIEFV 289
Db 312 LQNRPSPEIVERIITEAVDVEKYEPTDVLPSVLGLNCLNLCMCPYIEFL 359

RESULT 10
US-08-905-223-358
; Sequence 358, Application US/08905223
; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505

QY 272 VGLIGNCILMKQYIEFVADRLLVELGFSKVFOAENPDPFPMENISLEGKTNPFKEKRVSEY 331
Db 243 CDLIGNSLMSQTIIEFVADRLLLECLGCKSIKPHSKNPFNMWDKISLQKTNPFKEKRVADY 302
QY 332 QRPVMAETTDNVFTLDADP 351
Db 303 QKSGVNAQRKDHVFLCLNTEF 322

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Win95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/905,223
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 358:
SEQUENCE CHARACTERISTICS:
LENGTH: 147 amino acids
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
ORIGINAL SOURCE:
ORGANISM: Homo Sapiens
TISSUE TYPE: Brain
FEATURE:
NAME/KEY: sig_peptide
LOCATION: -101..-1
IDENTIFICATION METHOD: Von Heijne matrix
OTHER INFORMATION: score 4.4
OTHER INFORMATION: seq ISHILAFFAASDG/IV
US-08-905-223-358

Query Match 42.2%; Score 768; DB 3; Length 147;
Best Local Similarity 100.0%; Pred. No. 2.5e-75;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIPPPIQYPDIMWKYKQAQ 60
Db 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIPPPIQYPDIMWKYKQAQ 60
QY 61 ASFTWAEVDLSKDLPHWNKLKADEKYFISHILAFFAASDGI VNNENLVRFSEVQVPEA 120
Db 61 ASFTWAEVDLSKDLPHWNKLKADEKYFISHILAFFAASDGI VNNENLVRFSEVQVPEA 120
QY 121 RCYFGQILLENVHSEMYSLIIDTYIR 147
Db 121 RCYFGQILLENVHSEMYSLIIDTYIR 147

RESULT 11
US-09-270-767-46236
; Sequence 46236, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 46236
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46236

Query Match 32.9%; Score 599; DB 4; Length 149;
Best Local Similarity 76.5%; Pred. No. 5.7e-57;
Matches 114; Conservative 14; Mismatches 21; Indels 0; Gaps 0;

QY 196 VFFSGSFAAIFWLKKGRLMPLGLTFSGNELISRDGLHCDPACILMFOYLVNKPSEERREI 255
```


Db 1 IFFSGSFASIFWKKRGLMPLGTLFSNELISRDEGLHCDFAVLMMFQHLVQRPKRRIIEII 60
QY 256 VDAVKIEOEFLEALPVGLIGNCILMKOYIEFVADRLAVELGFSKVQFAENPPDFMEMI 315
Db 61 RDAVAIEOEFLEALPVGLIGNCILMKOYIEFVADRLAVELGFSKVQFAENPPDFMEMI 120
QY 316 SLEGTNPFKEKRVSEYQRFVAVNAETTDNV 344
Db 121 SLDGKTNPFKEKRVSEYQRFVAVNAETTDNV 149

RESULT 12

US-09-902-540-10006
; Sequence 10006, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10006
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10006

Query Match 27.2%; Score 495; DB 4; Length 323;
Best Local Similarity 35.0%; Pred. No. 3.6e-45;
Matches 113; Conservative 63; Mismatches 125; Indels 22; Gaps 6;

QY 49 YPDIWKMYKQAOASFWTAEEVDLSKDLPHW-NKLKADKDYFISHILAFPAASDGINVENL 107
Db 3 YPQFEMRYNATKNTMTVEEVDFTDLVLRSKMTDAERHLIHLVAFATGDSIVGNL 62
QY 108 VERFQEQVQPCRCYFGQIILNHYSEMYSLIIDYIRDPKKEFLNAETMPYVKK 167
Db 63 VLNLKHLNAPARMYLSQLVEALHVOFYLLTDTYVDPDAERAMAAVNDNIFSIOR 122
QY 168 KADWALRWI-----ADRKSTFGER-----VVAFA-AVEGVFFSGSFAAIFWLKRGGLMP 215
Db 123 KAQFCWKWMDSIQGLDTLQTKAHRQFLNLCIFAGCIEGLFFFAFAFVYFLRSKGLN 182
QY 216 GLTFSENELISRDEGLHCDFACLMFQY-----LVNKPSEERVREIIVDAVKIEQEFLE 268
Db 183 GLAAGTNWFRDESAHMAFESIQVARKEEPLDFAQNERDVVAMLEAVECETOFAQD 242
QY 269 ALPVGLIGNCILMKOYIEFVADRLAVELGFSKVQFAENPPDFMEMI SLEGTNPFKEKRV 328
Db 243 LLSGGVGLUSVQDMRGYLSVADQRLQMLQMTVPVGTGKNPLHPMDLQDVQELTNFERRV 302
QY 329 SEYQRFVAVNAETTDNVFTLDADF 351
Db 303 SSYQVAVGGAATDVV--LDAAF 323

RESULT 13

US-09-328-352-4714
; Sequence 4714, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4714
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4714

Query Match 21.0%; Score 381.5; DB 4; Length 445;
Best Local Similarity 27.8%; Pred. No. 1.2e-32;
Matches 97; Conservative 76; Mismatches 143; Indels 33; Gaps 10;

QY 9 AAGLDQDERSSDTNESEIKSNEEPLRKSSRRFVIFPIQYPIWKMYKQAOASFWTAEE 68
Db 98 APGLELENGA-----QRQVDDKAMINCRADNLQLVPFKYEWAWQKYLDCGANHMPQE 152
QY 69 VLSKDLPHW---NKLKADKDYFISHILAFPAASDGINVENLVERFQEQVQPCRCYFG 125
Db 153 VNMNHDIALWKSENGLTEDERTIVMRSLGFFSTADSLVANNLVLAIRHITNPCEQOYL 212
QY 126 FOILLIENVHSEMYSLIIDYIRDPKKEFLNAETMPYVKKKADWALRW---IADRKST 182
Db 213 RQAFEEAIHTHAYQCYIESLGMDEGE---VFNMYREIPSVARKAANGKLYTQSUSDPTFH 269
QY 183 FG-----ERVAFAAV-EGVFFSGSFAAIFWLKRGGLMPGLTFSNELISRDEGLHC 232
Db 270 TGTPENDQRLRLNLIAPYCVLEGIFFYCGFTQILSMGRNKGNGVAFQFOYLIRDESML 329
QY 233 DPACLMP-QYLVNKP-----EERVREIIVDAVKIEQEFLEALPVGLIGNCILMKOY 285
Db 330 NFGIDMINQIKIENPHLWTAERFQEQEVIQMLLEGTMLEIYARDTMPRGVLGNASMMBEY 389
QY 286 IEFVADRLAVELGFSKVQ- AENPPDFM-ENISLEGTNPFKEKRVSEYQ 332
Db 390 LKFCNRRUSQLGLPEQFAGVTNPPFAMWSEMDLRKKNPFETRVTDYQ 438

RESULT 14

US-09-252-991A-19358
; Sequence 19358, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19358
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19358

Query Match 20.7%; Score 377.5; DB 4; Length 476;
Best Local Similarity 26.7%; Pred. No. 3.7e-32;
Matches 100; Conservative 72; Mismatches 154; Indels 49; Gaps 11;

QY 2 GDEPERAAAGLDQDE---RSSDNTNESEIKSNEEPLRKSSRRFV-----I 44
Db 99 GSVEEARAVSADSDAVARAKKALNDLDIQEGDLDEGSAARVQVQDKQMINARADNLQ 158
QY 45 FPIQYPIWKMYKQAOASFWTAEEVDLSKDLPHWNK---LKADEKYFISHILAFPAASDG 101
Db 159 VPPKYDWAWQKYLDCGANHMPQEVNMNADIALWKSDGLSEHERRIVMRNLGFFSTADS 218
QY 102 IVNENLVERFSQVQVPEARCFYGFQILLIENVHSEMYSLIIDYIRDPKKEFLNAET 161

```
Db      219 LVANNLVLAAYRLITNPECRQVILRQAFEEAIHTHAYQYCIESLGMDEGE---IFNNYHE 275
QY      162 MPYVKKKADWALRWI-----ADRKSTGERVVAFAAV-EGVFFSGSFAAIF 206
Db      276 IPSVAKKASWGLKYRFSISDPMPQTGTPETDRQ--FLRLIAYYCVLEGIFCYCGTQIL 333
QY      207 WLKKGGLMPLGFTSNELISDRDEGLHCDPAC-LMFQYLVNKP-----SEERVREIIVDAV 259
Db      334 SMGRNKMGTGTAEQFQYILRDESMHLNFGIDVINQIKIENPHLWDAQMKDEATQMLQGT 393
QY      260 KTEQEPLEALPVGLIGMNCILMKQYIEFVADRLLVELGFSKVQ-AENPPDFMENI-SL 317
Db      394 QLEIEYARDTMPRGVLGNNAAMMEDYLKFIANRLTQIGLKEEYPGTTNPPFPMMSIMDL 453
QY      318 EGTKNPFPEKRVSEYQ 332
Db      454 KKEKNFFETRVIEYQ 468

RESULT 15
US-09-198-452A-1059
; Sequence 1059, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffaix, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1059
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-1059

Query Match      18.4%; Score 335; DB 4; Length 365;
Best Local Similarity 26.6%; Pred. No. 1e-27;
Matches      84; Conservative      64; Mismatches 134; Indels      34; Gaps      6;

QY      44 IPIQYDIDWKYKQAQAFWTAEEVDLSKDLPHM--NKLKADKXFISHILAFPAASDG 101
Db      50 LVPIKKAWEHYLNGCANNWLPTVPWARDLWKSDESDERRVILLNLGFFSTAES 109
QY      102 IVNENLVERFSQEVQPEARCFYQILTIENVHSEMYSLIDTYIRDPKKBFLFNAIET 161
Db      110 LVGNIVLAIPIKHTNPEARQYLLRQAFEEAVHTHTFLYICESLGLDEGE---VFNAYNE 166
QY      162 MPYVKKKADWALRWIAD-----RKSTGERVVAFAAVEGVFFSGSPAA 204
Db      167 RASIRAKDDFQMTLVTDVLDPNFVSQSSBGLQGFKNLVGVYII---MEGIFFYSGFWM 222
QY      205 IPWLKKGMLPGLTFSNELISDEGLHCDPACLMFQYLVNKPSE-----ERVREIIVD 257
Db      223 ILSFRQNMGTGIGQYQYILRDETHLNFGLDINGIKENPEVMTTELQEEIIVALTEK 282
QY      258 AVKIQEFLTEALPVGLIGMNCILMKQYIEFVADRLLVELGFSKVQAEENPFDFM-ENIS 316
Db      283 AVELIEYAKDCLPGILGRSSMFDIVRHIADRLRLRIGLKPIYHSRNPFPWMSETMD 342
QY      317 LEGKTNPFPEKRVSEYQ 332
Db      343 LNKEKNFFETRVIEYQ 358
```

Search completed: October 30, 2005, 06:58:42
Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2005, 06:57:41 ; Search time 72 Seconds
(without alignments)
2037.960 Million cell updates/sec

Title: US-10-698-228-1
Perfect score: 1821
Sequence: 1 MGDPERPEAAGLDQDRSSS.....QRFVMAETDNTVFTLDADF 351

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1865214 seqs, 418043040 residues
Total number of hits satisfying chosen parameters: 1865214

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1821	100.0	351	15	US-10-698-228-1
2	1454	79.8	389	15	US-10-698-228-5
3	1454	79.8	389	16	US-10-408-765A-559
4	1454	79.8	389	16	US-10-733-878-457
5	1454	79.8	389	18	US-10-756-149-4899
6	1454	79.8	413	9	US-09-925-301-1347
7	1454	79.8	453	18	US-10-220-335-686
8	1454	79.8	453	18	US-10-450-763-45646
9	1425.5	78.3	386	15	US-10-403-571-76
10	1348	74.0	393	20	US-11-097-143-19182
11	1242	68.2	430	15	US-10-369-493-4122
					Sequence 1, Appli
					Sequence 5, Appli
					Sequence 457, App
					Sequence 4899, Ap
					Sequence 1347, App
					Sequence 686, App
					Sequence 45646, A
					Sequence 76, Appl1
					Sequence 19182, A
					Sequence 4122, Ap

ALIGNMENTS

RESULT 1
US-10-698-228-1
; Sequence 1, Application US/10698228
; Publication No. US2004007253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 1
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-1

Query Match 100.0%; Score 1821; DB 15; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.7e-162;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGDPERPEAAGLDQDRSSSDTNESEIKSNEEPLLRKSSRRFVFPFQYDINWKYKQAQ 60
Db 1 MGDPERPEAAGLDQDRSSSDTNESEIKSNEEPLLRKSSRRFVFPFQYDINWKYKQAQ 60

Sequence 8245, Ap
Sequence 5686, Ap
Sequence 3245, Ap
Sequence 7537, Ap
Sequence 10365, A
Sequence 22170, A
Sequence 22565, A
Sequence 3145, Ap
Sequence 215866,
Sequence 215845,
Sequence 113592,
Sequence 45287, A
Sequence 247984,
Sequence 269685,
Sequence 183113,
Sequence 200125,
Sequence 6922, Ap
Sequence 342, App
Sequence 13401, A
Sequence 215843,
Sequence 21987, A
Sequence 215840,
Sequence 2195, Ap
Sequence 45645, A
Sequence 37219, A
Sequence 10564, A
Sequence 215870, A
Sequence 19363, A
Sequence 45286, A
Sequence 51357, A
Sequence 66252, A
Sequence 50118, A
Sequence 69814, A
Sequence 68043, A

QY 61 ASFWTAEEVDLSKDLPHWNKLKADKDYFISHILAFPAASDGIIVNENLVERFSQEVQVPEA 120
DB 61 ASFWTAEEVDLSKDLPHWNKLKADKDYFISHILAFPAASDGIIVNENLVERFSQEVQVPEA 120
QY 121 RCFYGFQIILNIENHVSMSYLLIDTYIRDPKREFFLNAETMPYVKKKADWALRWIADRK 180
DB 121 RCFYGFQIILNIENHVSMSYLLIDTYIRDPKREFFLNAETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFVAVGVEFGSSFAAIFWLKRGMLPGITFSNELISRDEGLHCDPACLMFQ 240
DB 181 STFGERVVAFVAVGVEFGSSFAAIFWLKRGMLPGITFSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSEERREIIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLILVELGFS 300
DB 241 YLVNKPSEERREIIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLILVELGFS 300
QY 301 KVFAENPPDFMENISLEGKTNPFKRVSEYQRFVMAETTDNVFTLDADF 351
DB 301 KVFAENPPDFMENISLEGKTNPFKRVSEYQRFVMAETTDNVFTLDADF 351

RESULT 2

US-10-698-228-5
; Sequence 5, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 5
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-5

Query Match 79.8%; Score 1454; DB 15; Length 389;
Best Local Similarity 83.8%; Pred. No. 6.6e-128;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY 31 EPLLRKSRFFVIFPIQYDWMKYKQASFWTAEEVDLSKDLPHWNKLKADKDYFIS 90
DB 69 DEPLLRNPRRFVIFPIEYHDIWMKYKAEASFWTAEEVDLSKDIQHWESLKPEERYFIS 128
QY 91 HILAFPAASDGIIVNENLVERFSQEVQVPEARCFYGFQIILNIENHVSMSYLLIDTYIRDPK 150
DB 129 HVLAFPAASDGIIVNENLVERFSQEVQITEARCFYGFQIAMENIHSEMSYLLIDTYIKDPK 188
QY 151 KREFLNAETMPYVKKKADWALRWIADRKSTGERVVAFAAAGEVFFSGSFAAIFWLKK 210
DB 189 EREFLNAETMPYVKKKADWALRWIADRKSTGERVVAFAAAGEVFFSGSFAAIFWLKK 248
QY 211 RGLMPGLTFSNELISRDEGLHCDPACLMFQYLVNKPSEERREIIIVDAVKIEQEFTEAL 270
DB 249 RGLMPGLTFSNELISRDEGLHCDPACLMFQYLVNKPSEERREIIIVDAVKIEQEFTEAL 308
QY 271 PVGLIGMNCILMKQYIEFVADRLILVELGFSKVFOAENPPDFMENISLEGKTNPFKRVSE 330
DB 309 PVKLGIMNCTLMKQYIEFVADRLILVELGFSKVFRVENPPDFMENISLEGKTNPFKRVSE 368

QY 331 YQFVAVMAETTDNVFTLDADF 351
DB 369 YQMGVMSSTPENSFTLDADF 389

RESULT 3

US-10-408-765A-559
; Sequence 559, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Watnick, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-559

Query Match 79.8%; Score 1454; DB 16; Length 389;
Best Local Similarity 83.8%; Pred. No. 6.6e-128;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY 31 EPLLRKSRFFVIFPIQYDWMKYKQASFWTAEEVDLSKDLPHWNKLKADKDYFIS 90
DB 69 DEPLLRNPRRFVIFPIEYHDIWMKYKAEASFWTAEEVDLSKDIQHWESLKPEERYFIS 128
QY 91 HILAFPAASDGIIVNENLVERFSQEVQVPEARCFYGFQIILNIENHVSMSYLLIDTYIRDPK 150
DB 129 HVLAFPAASDGIIVNENLVERFSQEVQITEARCFYGFQIAMENIHSEMSYLLIDTYIKDPK 188
QY 151 KREFLNAETMPYVKKKADWALRWIADRKSTGERVVAFAAAGEVFFSGSFAAIFWLKK 210
DB 189 EREFLNAETMPYVKKKADWALRWIADRKSTGERVVAFAAAGEVFFSGSFAAIFWLKK 248
QY 211 RGLMPGLTFSNELISRDEGLHCDPACLMFQYLVNKPSEERREIIIVDAVKIEQEFTEAL 270
DB 249 RGLMPGLTFSNELISRDEGLHCDPACLMFQYLVNKPSEERREIIIVDAVKIEQEFTEAL 308
QY 271 PVGLIGMNCILMKQYIEFVADRLILVELGFSKVFOAENPPDFMENISLEGKTNPFKRVSE 330
DB 309 PVKLGIMNCTLMKQYIEFVADRLILVELGFSKVFRVENPPDFMENISLEGKTNPFKRVSE 368
QY 331 YQFVAVMAETTDNVFTLDADF 351
DB 369 YQMGVMSSTPENSFTLDADF 389

RESULT 4

US-10-733-878-457
; Sequence 457, Application US/10733878
; Publication No. US20040224408A1
; GENERAL INFORMATION:
; APPLICANT: Jean-Philippe Girard
; APPLICANT: Francois Amalric
; APPLICANT: Myriam Roussigne
; APPLICANT: Thomas Clouaire
; TITLE OF INVENTION: THAP PROTEINS AS NUCLEAR RECEPTORS FOR
; TITLE OF INVENTION: CHEMOKINES AND ROLES IN TRANSCRIPTIONAL REGULATION, CELL
; FILE REFERENCE: BIOBANK.012A
; CURRENT APPLICATION NUMBER: US/10/733,878

```

; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 60/432699
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/485027
; PRIOR FILING DATE: 2003-07-03
; NUMBER OF SEQ ID NOS: 535
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 457
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-733-878-457

Query Match
Best Local Similarity 79.8%; Score 1454; DB 16; Length 389;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY 31 BEPLLRKSRFVIFPIQYPDIMWKYKQAQSFWTAEVDLSKDLPHNKLKADKEYFIS 90
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 91 HILAFPAASDGIENLVRFSGQVQVPEARCFYGFQIILNVHSEMYSLIDTYIRDPK 150
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 129 HVLAFAASDGIENLVRFSGQVQVTEARCFYGFQIAMIHSEMYSLIDTYIKDPK 188
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 151 KREFLFNAIETMPYVKKKADWALRWIADRKSTFGERVVAFVAVGVFFSGSFAAIFWLKK 210
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 189 EREFLFNAIETMPYVKKKADWALRWIGDKKEATYGERVAVFAAVGIFFGSFAIFWLKK 248
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 211 RGLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEERVREIIVDAVKIEQEFLEAL 270
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 249 RGLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEERVREIIVDAVKIEQEFLEAL 308
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 271 PVGLIGNCILMKQYIEFVADRLVLELGFQVFAENPDPFPMENISLEGKTNFFPKRVSE 330
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 309 PVKLIGNMCTLMKQYIEFVADRLMLLELGFQVFRVENPDPFPMENISLEGKTNFFPKRVGE 368
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 331 YORFVAVMAETTDNVFTLDADF 351
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 369 YQRMGVMSPTENSFTLDADF 389
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||

RESULT 5
US-10-756-149-4899
; Sequence 4899, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; FILE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4899
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-756-149-4899

Query Match
Best Local Similarity 79.8%; Score 1454; DB 18; Length 389;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY 31 BEPLLRKSRFVIFPIQYPDIMWKYKQAQSFWTAEVDLSKDLPHNKLKADKEYFIS 90
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 69 DEPLLRNPRRFVIFPIEYHDIMQYKKAESFWTAEVDLSKDIQHWESLKPPEERYFIS 128
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 91 HILAFPAASDGIENLVRFSGQVQVPEARCFYGFQIILNVHSEMYSLIDTYIRDPK 150
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 129 HVLAFAASDGIENLVRFSGQVQVTEARCFYGFQIAMIHSEMYSLIDTYIKDPK 188
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 151 KREFLFNAIETMPYVKKKADWALRWIADRKSTFGERVVAFVAVGVFFSGSFAAIFWLKK 210
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 189 EREFLFNAIETMPYVKKKADWALRWIGDKKEATYGERVAVFAAVGIFFGSFAIFWLKK 248
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 211 RGLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEERVREIIVDAVKIEQEFLEAL 270
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 249 RGLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEERVREIIVDAVKIEQEFLEAL 308
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 271 PVGLIGNCILMKQYIEFVADRLVLELGFQVFAENPDPFPMENISLEGKTNFFPKRVSE 330
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 309 PVKLIGNMCTLMKQYIEFVADRLMLLELGFQVFRVENPDPFPMENISLEGKTNFFPKRVGE 368
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 331 YORFVAVMAETTDNVFTLDADF 351
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 369 YQRMGVMSPTENSFTLDADF 389
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||

RESULT 6
US-09-925-301-1347
; Sequence 1347, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1347
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1347

Query Match
Best Local Similarity 79.8%; Score 1454; DB 9; Length 413;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY 31 BEPLLRKSRFVIFPIQYPDIMWKYKQAQSFWTAEVDLSKDLPHNKLKADKEYFIS 90
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 93 DEPLLRNPRRFVIFPIEYHDIMQYKKAESFWTAEVDLSKDIQHWESLKPPEERYFIS 152
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 91 HILAFPAASDGIENLVRFSGQVQVPEARCFYGFQIILNVHSEMYSLIDTYIRDPK 150
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 153 HVLAFAASDGIENLVRFSGQVQVTEARCFYGFQIAMIHSEMYSLIDTYIKDPK 212
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 151 KREFLFNAIETMPYVKKKADWALRWIADRKSTFGERVVAFVAVGVFFSGSFAAIFWLKK 210
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 213 EREFLFNAIETMPYVKKKADWALRWIGDKKEATYGERVAVFAAVGIFFGSFAIFWLKK 272
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 211 RGLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEERVREIIVDAVKIEQEFLEAL 270
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 273 RGLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEERVREIIVDAVKIEQEFLEAL 332
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 271 PVGLIGNCILMKQYIEFVADRLVLELGFQVFAENPDPFPMENISLEGKTNFFPKRVSE 330
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 333 PVKLIGNMCTLMKQYIEFVADRLMLLELGFQVFRVENPDPFPMENISLEGKTNFFPKRVGE 392
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 331 YORFVAVMAETTDNVFTLDADF 351
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 393 YQRMGVMSPTENSFTLDADF 413
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||

RESULT 7
US-10-220-335-686
; Sequence 686, Application US/10220335
; Publication No. US20050175607A1
```

```

QY 151 KREFLFNAIETMPYVKKKADWALRWIADRKSTFGERVVAFVAVGVFFSGSFAAIFWLKK 210
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 189 EREFLFNAIETMPYVKKKADWALRWIGDKKEATYGERVAVFAAVGIFFGSFAIFWLKK 248
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 211 RGLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEERVREIIVDAVKIEQEFLEAL 270
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 249 RGLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEERVREIIVDAVKIEQEFLEAL 308
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 271 PVGLIGNCILMKQYIEFVADRLVLELGFQVFAENPDPFPMENISLEGKTNFFPKRVSE 330
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 309 PVKLIGNMCTLMKQYIEFVADRLMLLELGFQVFRVENPDPFPMENISLEGKTNFFPKRVGE 368
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 331 YORFVAVMAETTDNVFTLDADF 351
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 369 YQRMGVMSPTENSFTLDADF 389
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||

RESULT 6
US-09-925-301-1347
; Sequence 1347, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1347
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1347

Query Match
Best Local Similarity 79.8%; Score 1454; DB 9; Length 413;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY 31 BEPLLRKSRFVIFPIQYPDIMWKYKQAQSFWTAEVDLSKDLPHNKLKADKEYFIS 90
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 93 DEPLLRNPRRFVIFPIEYHDIMQYKKAESFWTAEVDLSKDIQHWESLKPPEERYFIS 152
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 91 HILAFPAASDGIENLVRFSGQVQVPEARCFYGFQIILNVHSEMYSLIDTYIRDPK 150
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 153 HVLAFAASDGIENLVRFSGQVQVTEARCFYGFQIAMIHSEMYSLIDTYIKDPK 212
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 151 KREFLFNAIETMPYVKKKADWALRWIADRKSTFGERVVAFVAVGVFFSGSFAAIFWLKK 210
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 213 EREFLFNAIETMPYVKKKADWALRWIGDKKEATYGERVAVFAAVGIFFGSFAIFWLKK 272
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 211 RGLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEERVREIIVDAVKIEQEFLEAL 270
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 273 RGLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEERVREIIVDAVKIEQEFLEAL 332
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 271 PVGLIGNCILMKQYIEFVADRLVLELGFQVFAENPDPFPMENISLEGKTNFFPKRVSE 330
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 333 PVKLIGNMCTLMKQYIEFVADRLMLLELGFQVFRVENPDPFPMENISLEGKTNFFPKRVGE 392
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 331 YORFVAVMAETTDNVFTLDADF 351
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY 393 YQRMGVMSPTENSFTLDADF 413
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (148)..(203)
; OTHER INFORMATION: Ribonucleotide reductase small subunit proteins domain
; OTHER INFORMATION: Identified by eMATRIX, accession number BL00368A, p-value=9e-174, Pfam score 36.98
; OTHER INFORMATION: 40, raw score of 36.98
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (167)..(440)
; OTHER INFORMATION: Ribonucleotide reductases domain identified by Pfam, accession name ribonuc_red, E-value=9e-174, Pfam score 45.646
; OTHER INFORMATION:
US-10-450-763-45646

Query Match          79.8%; Score 1454; DB 18; Length 453;
Best Local Similarity 83.8%; Pred. No. 8.2e-128;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0

Qy 31 EEPLLKSSRRRVFPPIQYPIWIKYKQAQASFTWTAEEVDLSKDLPHNKLKADEKIFYIS 90
Db 133 DEPLLRENPRRFVFPPIEHDIWQYKAEASFMTAAEEVDLSKDIQHWESLKPEERYFIS 192
Qy 91 HILAFFAASDGIWVNLVERSESQVPEARCFYGFQILLENVHSEMYSLLLDTYIRDPK 350
Db 193 HVLAFFAASDGIWVNLVERSESQVQTEARCFYGFQIAMENIHSEMYSLLLDTYIKDPK 252
Qy 151 KREFLFNAIETMPVYKKADWALRWIADRKSTFGERVVFAAFAAEGVFFSGSFAAIFWLKK 210
Db 253 EREFLFNAIETMPCVKKADWALRWIGDEATYGERVVAFAAEGIFFSGSFAAIFWLKK 312
Qy 211 RGLMPGIGTFSNELISRDEGLHCDFACLMFOYLVNKPSEERVREIIVDAVKIQEFLTEAL 470
Db 313 RGLMEGLTFSNELISRDEGLHCDFACLMFKHLVHKPSEERVREIINAVRIQEFLTEAL 372
Qy 271 PVGLIGHNCTLMKOYIEFVADRLIVELGFSKVFOAENPDPFDMENISLGKTNFFEKRVSE 330
Db 373 PVKLIGHNCTLMKOYIEFVADRLMLELGLFSKVFRVENPDPFDMENISLGKTNFFEKRVGE 432
Qy 331 YQRFVMAETTTDNNVFTLDADF 351
Db 433 YQRMGVMSPTNSFTLDADF 453

RESULT 9
US-10-403-571-76
; Sequence 76, Application US/10403571
; Publication No. US20040068763A1
; GENERAL INFORMATION:
; APPLICANT: Hopkins, Nancy
; APPLICANT: Golling, Gregory
; APPLICANT: Amsterdam, Adam
; APPLICANT: Sun, Zhaoxia
; TITLE OF INVENTION: Developmental Mutations in Zebrafish
; FILE REFERENCE: 01997/539002
; CURRENT APPLICATION NUMBER: US/10/403.571
; CURRENT FILING DATE: 2003-03-25
; PRIOR APPLICATION NUMBER: US 60/368,760
; PRIOR FILING DATE: 2002-03-29
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Danio rerio
US-10-403-571-76

Query Match          78.3%; Score 1425.5; DB 15; Length 386;
Best Local Similarity 78.8%; Pred. No. 3.1e-125;
Matches 268; Conservative 36; Mismatches 31; Indels 5; Gaps 1

Qy 17 RSSSDTWESEIKS-----NEEPLLRKSSRRRVFPPIQYPIWIKYKQAQASFTWTAEEVDL 71

```


Db 47 RKIFDESGQKAKGAVEEPLLENKPHRFVIFPIQYHDIWQYKKAASFWTAEEVDL 106
QY 72 SKDLPHNKLKADKEYFISHILAFPAASDGIWVNLVERFSEOVQVPEARCFYGFQIILE 131
Db 107 SKDLQWLSKDEERYFISHVLAFFAASDGIWVNLVERFSEOVQVTEARCFYGFQIAE 166
QY 132 NVHSEMYSLIIDTYIRDPKRRFLFNAIETMPYVKKADWALRWIADRKSTFGERVVAF 191
Db 167 NIHSEMYSLIIDTYIKDSKEREFLFNAIETMPCVKKADWALNWDGNARYGERVVAFA 226
QY 192 AVEGVFFSGSPAIAFWLKRGLMPGLTFSNELISRDEGLHCDPACLMFOYLVNKPSEERV 251
Db 227 AVEGVFFSGSPASIFWLKRGLMPGLTFSNELISRDEGLHCDPACLMFKHLINKPSEETV 286
QY 252 REIIVDAVKIEOEFTALPVLGIGNCILMKOYIEFVADRLLVLFVGFSGVQAEPPDF 311
Db 287 KKIINVAIEOEFTDALPVKLIGNCILMKOYIEFVADRLLVLFVGFSGVQAEPPDF 346
QY 312 MENISLEGKTNFFKRVSEYQRFVMAETTDNVFTILDADP 351
Db 347 MENISLEGKTNFFKRVSEYQRFVMAETTDNVFTILDADP 386

RESULT 10

US-11-097-143-19182
; Sequence 19182, Application US/11097143
; Publication No. US2005020858A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19182
; LENGTH: 393
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-19182

Query Match 74.0%; Score 1348; DB 20; Length 393;
Best Local Similarity 73.9%; Pred. No. 6.3e-118;
Matches 252; Conservative 41; Mismatches 48; Indels 0; Gaps 0;
QY 11 GLDQDERSSDTNESEIKSNEBPLKRRRFFVIFPIQYDPDWKMYKQACASFWTAEEVD 70
Db 53 GIGKANSIMKSVTFPPDSLEPLRENPRRFFVIFPIQYHDIWQYKKAASFWTAEEVD 112
QY 71 LSKDLPHNKLKADKEYFISHILAFPAASDGIWVNLVERFSEOVQVPEARCFYGFQILI 130
Db 113 LSKDLTDWHLKDDERHSHVLAFFAASDGIWVNLVERFSEOVQVTEARCFYGFQIAM 172
QY 131 ENVHSEMYSLIIDTYIRDPKRRFLFNAIETMPYVKKADWALRWIADRKSTFGERVVAF 190

Db 173 ENVHSEMYSLIIDTYIRDPKRRFLFNAIETMPYVKKADWALSWISSKSAFGERIIAF 232
QY 191 AAVEGVFFSGSPAIAFWLKRGLMPGLTFSNELISRDEGLHCDPACLMFOYLVNKPSEER 250
Db 233 AAVEGVFFSGSPASIFWLKRGLMPGLTFSNELISRDEGLHCDPACLMFOHLVQRPKRER 292
QY 251 VREIIVDAVKIEOEFTALPVLGIGNCILMKOYIEFVADRLLVLFVGFSGVQAEPPDF 310
Db 293 IIEIIVDAVAIEOEFTDALPVNLIGNCILMKOYIEFVADRLLVLFVGFSGVQAEPPDF 352
QY 311 FMENISLEGKTNFFKRVSEYQRFVMAETTDNVFTILDADP 351
Db 353 FMENISLEGKTNFFKRVSEYQRFVMAETTDNVFTILDADP 393

RESULT 11

US-10-369-493-4122
; Sequence 4122, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 4122
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Neurospora crassa
US-10-369-493-4122

Query Match 68.2%; Score 1242; DB 15; Length 430;
Best Local Similarity 62.6%; Pred. No. 6.7e-108;
Matches 243; Conservative 52; Mismatches 47; Indels 46; Gaps 6;
QY 10 AGLDQDERSSDTNESEIKSNEBPLKRRRFFVIFPIQYDPDWKMYKQACASFWTAEEVDLSKDLPHW-NKLKADKEYFISHILAF 96
Db 43 AEIDAENANKKAAEAKQMAPTLKPEANEPLLTENPQRFVLPFIKYHEVCHASQTQTT 102
QY 52 -----IWKMYKQACASFWTAEEVDLSKDLPHW-NKLKADKEYFISHILAF 96
Db 103 RKVVHGLIGYPHYRIWQYKKAASFWTAEEVDLSKDLHWNRLNDEKKEFFISHILAF 162
QY 97 AASDGIWVNLVERFSEOVQVPEARCFYGFQIILENVHSEMYSLIIDTYIRDPKRRFLF 156
Db 163 AASDGIWVNLVERFSEOVQVPEARCFYGFQIILENVHSEMYSLIIDTYIRDPKRRFLF 222
QY 157 NAIETMPYVKKADWALRWIADRKSTFGERVVAFVAAVEGVFFSGSPAIAFWLKRGLMPG 216
Db 223 NAIDTIPICRKADWALRWITDKSSTFAQRILVAFVAAVEGVFFSGSPAIAFWLKRGLMPG 282
QY 217 LTFSENLISRDEGLHCDPACLMFOYLVNKPSEERVRIIIVDAVKIEOEFTALPVLGIL 276
Db 283 LTFSENLISRDEGLHCDPACLMFOYLVNKPSEERVRIIIVDAVKIEOEFTALPVLGIL 342
QY 277 MNCILMKOYIEFVADRLLVLFVGFSGVQAEPPDFMENISLEGKTNFFKRVSEYQRFVAF 336
Db 343 MNAIDLMKQYIEFVADRLLVLFVGFSGVQAEPPDFMENISLEGKTNFFKRVSEYQRFVAF 402
QY 337 M-----AETTDNV-----FTILDADP 351
Db 403 MNSTKKAADAAEVAKNENGSGDTFTDEDF 430

RESULT 12

US-10-128-714-8245
; Sequence 8245, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in *Aspergillus fumigatus* and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; PRIOR FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8245
; LENGTH: 405
; TYPE: PRT
; ORGANISM: *Aspergillus fumigatus*
US-10-128-714-8245

Query Match 67.5%; Score 1230; DB 14; Length 405;

Best Local Similarity 67.0%; Pred. No. 8.2e-107;
Matches 234; Conservative 52; Mismatches 49; Indels 14; Gaps 3;
QY 16 EBSDDTNESEIKSNEEPLLRKSSRRFVFPPIQYDIPWYKQAQASFWTAEEVDLSKDL 75
DB 58 EASKVAQSKELEAN-EPILQENPHRFVLPPIKYHEIWQYKAEASFWTAEEIDLSKDL 116
QY 76 PHW-NKLKADKYFISHILAFPAASDGIYVNEVLFSEVQVPEARCFYGFQILLNVH 134
DB 117 HDNNELNDDERYFTHVLAFAASDGIYVNEVLFSEVQVPEARCFYGFQIMLENH 176
QY 135 SEMYSLLIDTYIRDPKRRBFLNAIETMPYVKKADWALRWIADRKSTTGERVVAFAVE 194
DB 177 SETYSLLIDTYIKEPKORTYLFDAIDTVPCIRKKAQWAKRWISDKBSTFAQRLVAFAAVE 236
QY 195 GVFFSGSFAAIFWLKKGMLPGLTFSENLISRDGLHCDFACLMFOYLVNKPSEERVREI 254
DB 237 GIFFSGSFASIFWLKKGMLPGLTFSENLISRDGLHCDFACLMFOYLVNKPSEERVREI 296
QY 255 IYDAVKIEOEFLTEALPVGLIGNCTLMKQYIEFVADRLLVGLVSKVFOAENPPDFMEN 314
DB 297 IVEAVAIEQEFITDALPVALLGNKSLMKQYIEFVADRLLVGLVSKVFOAENPPDFMES 356
QY 315 ISLEGKTNFPEKRVSEYQRFVMAET-----TDNVFTLDADF 351
DB 357 ISLAGKTNFPEKRVSEYQRFVMAETKQASQDTAKDANDGGGLCPDED 405

RESULT 13

US-10-369-493-5686
; Sequence 5686, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5686
; LENGTH: 381
; TYPE: PRT
; ORGANISM: *Caenorhabditis elegans*
US-10-369-493-5686

Query Match 66.7%; Score 1214; DB 15; Length 381;

Best Local Similarity 67.2%; Pred. No. 2.4e-105;
Matches 231; Conservative 48; Mismatches 61; Indels 4; Gaps 2;

QY 12 LDQERSSD--TNESEIK--SNEEPLLRKSSRRFVFPPIQYDIPWYKQAQASFWTAEE 67
DB 38 VDQTKAASAEETNNESEVNELDADPEMLQDLNRFVIFPLKHHDINWYKXAVASFWTVE 97
QY 68 EVDLSKDLPHWNKLGKADKYFISHILAFPAASDGIYVNEVLFSEVQVPEARCFYGFQ 127
DB 98 EVDLGKDMNDWEKMGDEQYFISILAFPAASDGIYVNEVLFSEVQVSEARFFYGFQ 157
QY 128 ILIENVHSEMYSLIDTYIRDPKRRBFLNAIETMPYVKKADWALRWIADRKSTTGERV 187
DB 158 IAIENHSEMYSKLIETIIRDETERTLNFNAVDEFFIKKKADWALRWISDKKSAFARL 217
QY 188 VAPAAVEGVFFSGSFAAIFWLKKGMLPGLTFSENLISRDGLHCDFACLMFOYLVNKP 247
DB 218 IAFAAVEGIFSGSFASIFWLKKGMLPGLTFSENLISRDGLHCDFACLMFOYLVNKP 277
QY 248 EERVREIIVDAVKIEOEFLTEALPVGLIGNCTLMKQYIEFVADRLLVGLVSKVFOAEN 307
DB 278 QQRIYDIIDKDAVAIEQEFITDALPVALLGNKSLMKQYIEFVADRLLVGLVSKVFOAEN 337
QY 308 PPDFMENISLEGKTNFPEKRVSEYQRFVMAETTDNVFTLDADF 351
DB 338 PPDFMENISDGKTNFPEKRVSEYQRFVMAETTDNVFTLDADF 381

RESULT 14

US-10-128-714-3245
; Sequence 3245, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in *Aspergillus fumigatus* and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; PRIOR FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3245
; LENGTH: 381

```
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-3245

Query Match          65.5%; Score 1192.5; DB 14; Length 381;
Best Local Similarity 65.6%; Pred. No. 2.6e-103;
Matches 229; Conservative 51; Mismatches 50; Indels 19; Gaps 4;

QY 16 ERSSDNESEIKSNEBRLKSSRRFVFPPIQYDPDIWKMYKQAOASFWTAEVDLSKDL 75
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 39 EASKVAQSIKELEAM-EPULQENPHRFVFPPIKTHIWMYKKAASFWTAEIDLSKDL 97
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 76 PHW-NKLKADKYPFISHILAFPAASDGI VNNELVRFSEVQVPEARCFYGFQIILIENH 134
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 98 HDNNRLNDDERYFISHVLAFPAASDGI VNNF-----NEVQVPEARCFYGFQIIMIENH 152
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 135 SEMYSLIDITYTRDPKRBFLFNAIETMPYVKKADWALRWIADRKSTFGERVVAFVAE 194
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 153 SETYSLIDITYKEPKORTYLPDAITPICIRKKAQWAMRWISDKESTTFAQLVAFVAE 212
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 195 GVFFSGSPAIFWLKKGRLMPLGTFNSNELISRDEGLHCDPACLMFOYLVNKPSEERVREI 254
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 213 GIFFSGSFASIFWLKKGRLMPLGTFNSNELISRDEGLHCDPACLLFSLNHRPDKKVIEDI 272
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 255 IVDAVKIEQEFILTEALPVGLIGMNCILMKQYIEFVADRLILVELGFSKVFOAENPDPFMEN 314
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 273 IVEAVAIEQEFITDALPVALLGNNSKLMQYIEFVADRLILVALGNKYFNSPNPFDFMES 332
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 315 ISLEGKTNFFKRVSEYQRFVAVMAET-----TDNVFTLDADF 351
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 333 ISLAGKTNFFKRVGYQKAGVMASTKQDASQDTAKDANDGGLCFDEDF 381
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 15
US-10-032-585-7537
; Sequence 7537, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7537
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7537

Query Match          65.5%; Score 1192; DB 14; Length 413;
Best Local Similarity 63.7%; Pred. No. 3.2e-103;
Matches 225; Conservative 58; Mismatches 64; Indels 6; Gaps 3;

QY 5 ERPEAAGLDQDRSSSDTNESEIK--SNEEPLLRKSSRRFVFPPIQYDPDIWKMYKQAOAS 62
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 EKDDSLKXHQDPLAKXKHVRHKLQLEAEPLLVENKRRYVMPPIRYHINWPKKAEAS 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 63 FWTAEEVDLSKDLPHW-NKLKADKYPFISHILAFPAASDGI VNNELVRFSEVQVPEAR 121
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 FWTAEEIDLSKDLDDWNNKLNENERYFISKVLAFPAASDGI VNGENLIENFSTEVQLPEAK 180
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 122 CEVGFQIILIENHSEMYSLIDITYTRDPKRBFLFNAIETMPYVKKADWALRWIADRS 181
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 SFYGFQIMWNIHSETYSLIDITYKDPQEAADYLFNAIANIPICIQKADWAIKWIQDDEA 240
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 182 TFGERVVAFVAEVEGFFSGSPAIFWLKKGRLMPLGTFNSNELISRDEGLHCDPACLMFOY 241
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 LYGERLVAFVAEVEGFFSGSPAIFWLKKGRLMPLGTFNSNELISRDEGLHCDPACULFAH 300
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
QY 242 LVNKPSEERVREIIVDAVKIEQEFILTEALPVGLIGMNCILMKQYIEFVADRLILVELGFSK 301
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 LQNRPSPEIVERIITEAVDVEKEYFTDALPVSLIGMNCILMKQYIEFVADRLILVALGNKK 360
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 302 VFOAENPDPFMENISLEGKTNFFKRVSEYQRFVAVMAETD---NVFTLDADF 351
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 361 VYNTNPPDFMENISLAGKTNFFKRVSDYQKAGVMAKTENKEADAFTFDEDF 413
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

Search completed: October 30, 2005, 07:04:05
Job time : 73 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 30, 2005, 06:58:47 ; Search time 3469 Seconds

(without alignments)
4902.791 Million cell updates/sec

Title: US-10-698-228-1

Perfect score: 1821

Sequence: 1 MGDPERPEAAGLDQDRSSS.....QRPVMAETTDNVFTLDADF 351

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10698228/runat_26102005_100714_5107/app_query.fasta.1.519
-DB=GenEmbl -QFWT=fastcap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228@cgn_1_1_2527@runat_26102005_100714_5107 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOB=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_hg.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	100.0	1053	6	AR454868
2	1821	100.0	1053	6	AR454877
3	1821	100.0	1053	6	BD064764
4	1821	100.0	1053	6	BD064773

5	1821	100.0	1053	6	BD093077	Novel pro
6	1821	100.0	1053	6	BD093086	Novel pro
7	1821	100.0	1056	6	CQ714252	Sequence
8	1821	100.0	1056	9	AB166671	Homo sapi
9	1821	100.0	1081	6	AR454870	Sequence
10	1821	100.0	1081	6	BD064766	Novel pro
11	1821	100.0	1081	6	BD093079	Novel pro
12	1821	100.0	4955	6	AR454869	Sequence
13	1821	100.0	4955	6	BD064765	Novel pro
14	1821	100.0	4955	6	BD093078	Novel pro
15	1821	100.0	4955	9	AB036063	Homo sapi
16	1817	99.8	1601	6	BD156916	Primer fo
17	1817	99.8	1601	6	AX877905	Sequence
18	1817	99.8	1601	9	AK001965	Homo sapi
19	1725	94.7	4532	10	BC058103	Mus muscu
20	1627.5	89.4	171737	10	AC122379	Mus muscu
21	1514	83.1	900	9	AB163437	Homo sapi
22	1474	80.9	4571	9	BC042468	Homo sapi
23	1460.5	80.2	1649	9	BC001886	Homo sapi
24	1460.5	80.2	1653	9	BC030154	Homo sapi
25	1460.5	80.2	1794	9	S40301	Homo sapien
26	1460.5	80.2	2500	6	AR300103	Sequence
27	1460.5	80.2	2500	6	AR380825	Sequence
28	1460.5	80.2	2500	6	AR438503	Sequence
29	1460.5	80.2	2500	6	AR562011	Sequence
30	1460.5	80.2	2500	6	AX333242	Sequence
31	1460.5	80.2	2500	6	AX333687	Sequence
32	1460.5	80.2	2500	6	AX334345	Sequence
33	1460.5	80.2	2500	9	HSRR2SS	H.sapiens R
34	1456.5	80.0	3039	9	AK092671	Homo sapi
35	1445.5	79.4	3203	9	AK123010	Homo sapi
36	1444	79.3	1538	5	CR762018	Xenopus t
37	1444	79.3	1611	5	CR848103	Xenopus t
38	1433	79.2	2111	10	MUSRNRM2A	Mouse ribon
39	1439	79.0	1406	5	CR353642	Gallus ga
40	1439	79.0	1647	5	BC080161	Xenopus t
41	1438	79.0	217774	2	AC120603	Rattus no
42	1431	78.6	1605	5	BC047975	Xenopus l
43	1429	78.5	1170	6	E14825	cDNA encodi
44	1428.5	78.4	1600	5	BC041209	Xenopus l
45	1425.5	78.3	1328	5	DRU57965	Danio rerio

ALIGNMENTS

RESULT 1	AR454868	Sequence 2 from patent US 6682917.	1053 bp	DNA	linear	PAT 20-FEB-2004
LOCUS	AR454868	Sequence 2 from patent US 6682917.	1053 bp	DNA	linear	PAT 20-FEB-2004
DEFINITION	AR454868	Sequence 2 from patent US 6682917.	1053 bp	DNA	linear	PAT 20-FEB-2004
ACCESSION	AR454868	Sequence 2 from patent US 6682917.	1053 bp	DNA	linear	PAT 20-FEB-2004
VERSION	AR454868.1	GI:42688823	1053 bp	DNA	linear	PAT 20-FEB-2004
KEYWORDS	AR454868.1	GI:42688823	1053 bp	DNA	linear	PAT 20-FEB-2004
SOURCE	Unknown.	GI:42688823	1053 bp	DNA	linear	PAT 20-FEB-2004
ORGANISM	Unknown.	GI:42688823	1053 bp	DNA	linear	PAT 20-FEB-2004
REFERENCE	1 (bases 1 to 1053)	GI:42688823	1053 bp	DNA	linear	PAT 20-FEB-2004
AUTHORS	Nakamura, Y., Arakawa, H. and Tanaka, H.	GI:42688823	1053 bp	DNA	linear	PAT 20-FEB-2004
TITLE	Protein having a ribonucleotide Reductase activity and a DNA thereof	GI:42688823	1053 bp	DNA	linear	PAT 20-FEB-2004
JOURNAL	Patent: US 6682917-A 2 27-JAN-2004;	GI:42688823	1053 bp	DNA	linear	PAT 20-FEB-2004
FEATURES	Location/Qualifiers	GI:42688823	1053 bp	DNA	linear	PAT 20-FEB-2004
source	1. .1053	GI:42688823	1053 bp	DNA	linear	PAT 20-FEB-2004
ORIGIN	/organism="genomic DNA"	GI:42688823	1053 bp	DNA	linear	PAT 20-FEB-2004

Alignment Scores:	1.01e-170	Length:	1053
Pred. No.:	1821.00	Matches:	351
Score:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-698-228-1 (1-351) x AR454868 (1-1053)		Db	1021	ACAGATAACGTCCTTCCACCTTGGATGCAGATTTT	1053
RESULT 2		AR454877	AR454877	1053 bp	DNA
LOCUS		AR454877	Sequence 12 from patent US 6682917.	linear	PAT 20-FEB-2004
DEFINITION		AR454877	Sequence 12 from patent US 6682917.		
ACCESSION		AR454877	GI:42688832		
VERSION		AR454877.1	GI:42688832		
KEYWORDS		Unknown.			
SOURCE		Unknown.			
ORGANISM		Unclassified.			
REFERENCE		1 (bases 1 to 1053)			
AUTHORS		Nakamura, Y., Arakawa, H. and Tanaka, H.			
TITLE		Protein having a ribonucleotide Reductase activity and a DNA thereof			
JOURNAL		Patent: US 6682917-A 12 27-JAN-2004;			
FEATURES		Location/Qualifiers			
source		1..1053			
ORIGIN		/organism="unknown"			
		/mol_type="genomic DNA"			
Alignment Scores:		1.01e-170	Length:	1053	
Pred. No.:		1821.00	Matches:	351	
Score:		100.00%	Conservative:	0	
Percent Similarity:		100.00%	Mismatches:	0	
Best Local Similarity:		100.00%	Indels:	0	
Query Match:		100.00%	Gaps:	0	
DB:		6			
US-10-698-228-1 (1-351) x AR454877 (1-1053)		Qy	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer	20
		Db	1	ATGGCGCACCCGGAAGCGCGGCGCGGCTGGATCAGATGAGATCATCTTCA	60
		Qy	21	AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg	40
		Db	61	GACACCAACAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC	120
		Qy	41	ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln	60
		Db	121	CGGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAATGTATAACAGGCACAG	180
		Qy	61	AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrPheSer	80
		Db	181	GCITTCCTCTGGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAACAAG	240
		Qy	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp	100
		Db	241	CTTAAAGCAGATGAGAGTACTTCTCATCTCACATCTTAGCCTTTTTCGAGCCAGTGAT	300
		Qy	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla	120
		Db	301	GGATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGCTCCAGAGGT	360
		Qy	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140
		Db	361	CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTG	420
		Qy	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160
		Db	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA	480
		Qy	161	ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys	180
		Db	481	ACATGCCCTTATGTTAAGAAAAAGCAGATTTGGCCTTCGATGATAGCAGATAGAAA	540
		Qy	181	SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly	200
		Db	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGTCTTCTTCTCAGGA	600
		Qy	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer	20
		Db	1	ATGGCGCACCCGGAAGCGCGGCGCGGCTGGATCAGATGAGATCATCTTCA	60
		Qy	21	AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg	40
		Db	61	GACACCAACAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC	120
		Qy	41	ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln	60
		Db	121	CGGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAATGTATAACAGGCACAG	180
		Qy	61	AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrPheSer	80
		Db	181	GCITTCCTCTGGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAACAAG	240
		Qy	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp	100
		Db	241	CTTAAAGCAGATGAGAGTACTTCTCATCTCACATCTTAGCCTTTTTCGAGCCAGTGAT	300
		Qy	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla	120
		Db	301	GGATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGCTCCAGAGGT	360
		Qy	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140
		Db	361	CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTG	420
		Qy	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160
		Db	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA	480
		Qy	161	ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys	180
		Db	481	ACCATGCCCTTATGTTAAGAAAAAGCAGATTTGGCCTTCGATGATAGCAGATAGAAA	540
		Qy	181	SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly	200
		Db	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGTCTTCTTCTCAGGA	600
		Qy	201	SerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer	220
		Db	601	TCTTTTGTCTATATCTTGCTTAAAGAAAGAGAGGTCTTATGCGAGACTCACTTTTTC	660
		Qy	221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln	240
		Db	661	ANTGAACCTATCAGCAGATGAGAGACTTCACTGTGACTTTTCTGCTGCTGATGTTCCAA	720
		Qy	241	TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys	260
		Db	721	TACTTAGTAAATAAGCCTTCAGAAAGAGGTGAGGAGATCATTTGTGATGCTGTCAAA	780
		Qy	261	IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle	280
		Db	781	ATTGAGCAGAGTATTAACAGAGCCTTCGCCAGTTGGCCTCATTTGGAATGGAATTCATT	840
		Qy	281	LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer	300
		Db	841	TTGATGAACACAGTACATGAGTTGTAGCTGACAGATTAATCTGTGAACTTGGATTTCTCA	900
		Qy	301	LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys	320
		Db	901	AAGGTTTTTTCAGCAGAAAAATCCTTTTGTATTTTATGMAAAACATTTCTTTAGAAGGAAA	960
		Qy	321	ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr	340
		Db	961	ACAAATTTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGTATTGCGAGAAACC	1020
		Qy	341	ThrAspAsnValPheThrLeuAspAlaAspPhe	351

```

QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGTCTATATCTTGGCTAAAGAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
QY 221 AnGluLeuLysSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACCTCATCAGCAGAGATGAAGACTTCATGTGACTTTGCTTGGCTGTATGTTCCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluLeuIleValAspAlaValLys 260
Db 721 TACTTAGTAATTAATAGCTTCAGAGAAAGGGTCAGGAGATCATGTTGATGCTGTCAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 781 ATTGAGCAGGAGTCTTTTAAACAGAGCCTTGGCCATTGGCTCATTTGGAATGATTCAT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 841 TTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGTGGAATTCGATTC 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 901 AAGGTTTTTCAGCGAGAAATCCTTTGATTTTATGGAACAATTTCTTTAGAGGAAAA 960
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 961 ACAATTTCTTTGAGNACAGTTTCAGAGTATCAGCGTTTGCATTTATGGCAGTAATGG 1020
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1021 ACAGATAACGCTTCCACTTGGATGAGATTTT 1053

RESULT 3
BD064764
LOCUS
DEFINITION
ACCESSION BD064764
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
PD 02-OCT-2001
PF 27-JUN-2000 JP 2000192401
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N15/09, A61K31/71, A61K38/00, A61K38/43, A61K45/00, A61K48/00, PC
A61P35/00,
PC A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC
PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53, (C12N15/09,
PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53, (C12N15/09,
PC C12N15/09, C12R1:91), (C12N9/02, C12R1:91), (C12P21/02, C12R1:91),
PC (C12P21/08, C12R1:91), (C12N15/00, A61K37/02, A61K37/48, C12N5/00,
PC (C12N15/00, C12R1:91), (C12N15/00, C12R1:19)
CC Novel protein and DNA thereof
FH Key Location/Qualifiers
FT source 1..1053
FT /organism='Homo sapiens (human)'
FT Location/Qualifiers
1..1053
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'

ORIGIN

```

```

Alignment Scores:
Pred. No.: 1,01e-170 Length: 1053
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-698-228-1 (1-351) x BD064764 (1-1053)
QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
Db 1 ATGGCGACCCCGGAAAGCGCGAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
Db 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Db 121 CGTTTGTCTATCTTTCCATCCAGTACCTGATATTTGGAAAATGTATAACAGGCACAG 180
QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db 181 GCTTCCTTCTGGACAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAspSerAsp 100
Db 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGTAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTCGAGGTTCCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 361 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAATGTTCTCACTCAGAGATGTACAGTTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 421 CTGATAGACACTTACATCAGATGCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 481 ACCATGCCCTATGTTAAGAAAAAGCAGATGGCCCTTGGCATGGATGAGATAGAGAAAA 540
QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTAGAGAGGAGTTTCTTCTCAGGA 600
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGTCTATATCTTGGCTAAAGAGAGAGGCTTATGCCAGGACTCACTTTTTC 660
QY 221 AnGluLeuLysSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACCTCATCAGCAGAGATGAAGACTTCATGTGACTTTGCTGCTCATGTTTCCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260
Db 721 TACTTAGTAATTAATAGCTTCAGAGAAAGGGTCAGGAGATCATGTTGATGCTGTCAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 781 ATTGAGCAGGAGTCTTTTAAACAGAGCCTTGGCCAGTTCGCTGCTCATGTTTCCAA 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 841 TTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGTGGAATTCGATTC 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 901 AAGGTTTTTCAGCGAGAAATCCTTTGATTTTATGGAACAATTTCTTTAGAGGAAAA 960

```

```

QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 961 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTGCGATTATGGCAGAAACC 1020

QY 341 ThrAsnValPheThrLeuAspAlaAspPhe 351
Db 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053

RESULT 4
BD064773
LOCUS BD064773 1053 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel protein and DNA thereof.
ACCESSION BD064773
VERSION BD064773.1 GI:22610376
KEYWORDS JP 2001269184-A/10.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Nakamura,Y., Arakawa,H. and Tanaka,H.
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
JOURNAL 1 (bases 1 to 1053)
COMMENT YUSUKE NAKAMURA,HIROFUMI ARAKAWA,HIROSHI TANAKA PC
OS JP 2001269184-A 10 02-OCT-2001;
OS YUSUKE NAKAMURA,TAKEDA CHEMICAL INDUSTRIES LTD
OS Homo sapiens (human)
PN JP 2001269184-A/10
PD 02-OCT-2001
PI 27-JUN-2000 JP 2000192401
PF YUSUKE NAKAMURA,HIROFUMI ARAKAWA,HIROSHI TANAKA PC
PC C12N15/09,A61K31/711,A61K38/00,A61K38/43,A61K45/00,A61K48/00,PC
PC A61P35/00,
PC A61P43/00,C07K16/32,C07K16/40,C12N1/15,C12N1/19,C12N1/21 PC
PC C12N5/10,C12N9/02,
PC C12P21/02,C12P21/08,G01N33/15,G01N33/50,G01N33/53//(C12N15/09,
PC C12R1:91),
PC (C12N15/09,C12R1:19),(C12N9/02,C12R1:91),(C12P21/02,C12R1:91),
PC (C12P21/08,C12R1:91),C12N15/00,A61K37/02,A61K37/48,C12N5/00,
PC (C12N15/00,C12R1:91),(C12N15/00,C12R1:19)
CC Novel protein and DNA thereof
FH Key Location/Qualifiers
FT source 1..1053
FT /organism='Homo sapiens (human)'.

FEATURES
source
1..1053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1,01e-170 Length: 1053
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-698-228-1 (1-351) x BD064773 (1-1053)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 1 ATGGGGACCCGGAAGCGGGAAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 60

QY 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
Db 61 GACACCAACGAAGTGAATGAATGAATCAATGAGAGCCACTCTCTAGAAGAGTTCTCG 120

QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysGlnAlaGln 60
Db 121 CGGTTTGTTCATCTTCCAAATCCAGTACCTGATATTTGGAATGTTATAAACAGGCACAG 180

QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80

```

```

Db 181 GCCTTCCTTCGGCAGCAGAAAGAGTTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAspArg 100
Db 241 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCTTTTTCGAGCCAGTAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 301 GGAATTGTAAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTCGAGGTTCCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTATTAATGCAATTGAA 480
QY 161 ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrPheAlaAspArgLys 180
Db 481 ACCATGCCCTATGTTAAGAAAAGAGATTTGGCCCTTGGATGGATAGCAGATAGAAA 540
QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGluValPhePheSerGly 200
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGATTTTCTTCTCAGGA 600
QY 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTGCTGCTATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGAGTCACTTTTTC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTGTCTGCTGATGTTTCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 721 TACTTAGTAATAAGCCTTCAGAAAGAGGTCAGGAGAGATCATTTGTTGATGCTGTCAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 781 ATTGAGCAGGAGTTTAAACAGAAAGCCTTGCAGTTCGCTCATTTGGAATGAATTCAT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 841 TTGATGAAAACAGTACATTTGAGTTGTAGCTGACAGATTTACTTTGGAACTTTGGATTTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 901 AAGGTTTTTCAGGCAGAAATCCTTTGATTTTATGGAAAACATTTCTTTAGAGGAAAA 960
QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGTTTATGCGAGAAACC 1020
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053

RESULT 5
BD093077
LOCUS BD093077 1053 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel protein and its DNA.
ACCESSION BD093077
VERSION BD093077.1 GI:22638665
KEYWORDS WO 0100799-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE Novel protein and its DNA

```


JOURNAL		Patent: WO 0100799-A 1 04-JAN-2001; TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA	
COMMENT	OS	Homo sapiens (human)	
	PN	WO 0100799-A/1	
	PD	04-JAN-2001	
	PF	27-JUN-2000 WO 2000JP004189	
	PR	28-JUN-1999 JP 99P 181131, 06-JUL-1999 JP 99P 192391 PR	
	PI	21-JAN-2000 JP 00P 017770	
	PI	YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC	
	C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC		
	PC	A61K48/00, C07K16/40, G01N33/53, C12Q1/26	
	CC		
	CC		
FEATURES	PH	Key	Location/Qualifiers.
	source	1..1053	
ORIGIN	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
Alignment Scores:			
Pred. No.:	1.01e-170	Length:	1053
Score:	1821.00	Matches:	351
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-10-698-228-1 (1-351) x BD093077 (1-1053)			
QY	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer	20
Db	1	ATGGGCGACCCGAAAGCGCGAAGCGCGCTGGATCAGATGAGAGATCATCTTCA	60
QY	21	AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg	40
Db	61	GACACCAACAAAGTGAATAAAGTCAATGAAGAGCCACTCTCAAGAAAGAGTTCTCGC	120
QY	41	ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln	60
Db	121	CGGTTTTGTCACTTTCCATCCAGTACCTCGATATTGGAAATGTATTAACAGGCACAG	180
QY	61	AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys	80
Db	181	GCTTCCTCTCGACAGCAGAGAGTTCGACTTATCAAGGATCTCCCTCACTGGAACAAG	240
QY	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp	100
Db	241	CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCGACCCAGTGAT	300
QY	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla	120
Db	301	CGAATTGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT	360
QY	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140
Db	361	CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTG	420
QY	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160
Db	421	CTGATAGACACTTATCATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA	480
QY	161	ThrMetProTyrValLysLysLysAlaAspTyrPalaLeuArgTyrIleAlaAspArgLys	180
Db	481	ACCATGCCCTATGTTAAGAAAAGCAGATTGGGCTTCGGATGATAGCAGATAGAAA	540
QY	181	SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly	200
Db	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAAGAGGAGTTTCTTCTCAGGA	600
QY	201	SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer	220

Db	601																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																										
----	-----	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

DB: 6 Gaps: 0
US-10-698-228-1 (1-351) x BD093086 (1-1053)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
DB 1 ATGGGCGACCCCGGAAGCGCGGAGCGCGCTGGATCAGATGAGAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
DB 61 GACACCAACCAAGTGAATAAGTCAATGAAGACCACTCTTAAGAAGAAGTTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
DB 121 CGGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG 180
QY 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
DB 181 GCTTCTCTTGACACGACAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaIleSerAsp 100
DB 241 CTTAAGCAGATGAGAGTACTTCATCTCTACATCTTAGCCCTTTTTCAGCCAGTGAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 301 CGAATTGTAAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGCTCCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 480
QY 161 ThrMetProTyrValLysLysAlaAspTyrPheAlaLeuArgTyrPheAlaAspArgLys 180
DB 481 ACCATGCCCTATGTAAAGAAAAGCAGATTTGGCCCTTCGGATGGATAGCAGATAGAAA 540
QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
DB 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTCGTCTGTAAGAGGAGTCTTCTCTCAGGA 600
QY 201 SerPheAlaAlaIlePheThrPheLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 661 AATGAACCTCATCAGCAGATGAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB 721 TACTTAGTAATAAGCTTTCAGAAAGAGGCTCAGGAGATCATCTTGTATGCTGTCAAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 781 ATTGACGAGGAGTTTTTAAACAGAGCTTTGCCAGTTGGCTCATTTGGAATGAATTCATT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 841 TTGATGAACAGTACATTTAGTTTGTAGCTGACAGATTTCTTGTGGAACCTTGGATTTCTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 901 AAGGTTTTTCAGGCAGAAAATCCTTTGATTTTATGAAAACATTTCTTTTAGAAGGAAA 960
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 961 ACAAAATTTCTTTGAGAAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

DB 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053

RESULT 7
CQ714252
LOCUS CQ714252 1056 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 186 from Patent WO02068579.
ACCESSION CQ714252
VERSION CQ714252.1 GI:42275109
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 186 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source Location/Qualifiers
1..1056
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 1,01e-170 Length: 1056
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-698-228-1 (1-351) x CQ714252 (1-1056)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
DB 1 ATGGGCGACCCCGGAAGCGCGGAGCGCGCTGGATCAGATGAGAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
DB 61 GACACCAACCAAGTGAATAAGTCAATGAAGACCACTCTTAAGAAGAAGTTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
DB 121 CGGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG 180
QY 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
DB 181 GCTTCTCTTGACACGACAGAGGTGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaIleSerAsp 100
DB 241 CTTAAGCAGATGAGAGTACTTCATCTCTACATCTTAGCCCTTTTTCAGCCAGTGAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 301 CGAATTGTAAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGCTCCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 480
QY 161 ThrMetProTyrValLysLysAlaAspTyrPheAlaLeuArgTyrPheAlaAspArgLys 180
DB 481 ACCATGCCCTATGTAAAGAAAAGCAGATTTGGCCCTTCGGATGGATAGCAGATAGAAA 540

```

181 SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 200
Db TCTACTTTTGGGAAGAAGAGTGTGGCTTGTCTGTAGAGAGGATTTCTTCTCAGGA 600

201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db TCTTTTGTCTGTATATCTGGCTAAAGACAGAGAGTCTTATGCCAGGACTCACTTTTCC 660

221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 720

241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260
Db TACTTAGTAATAAAGCTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGCACAA 780

261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db ATTGAGCAGGAGTCTTTTAAACAGAGCCTTGCCAGTTGGCTCATTTGAATGATTCATT 840

281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db TTGATGAACACGATACATCAGTTTGTAGCTGACAGATTAATTTGTGGAATTCGATTCTCA 900

301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db AAGGTTTTTTCAGCGAGAAATCCCTTTGATTTTATGGAATAACATTTCTTTAGAGGAAAA 960

321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db ACAATTTCTTTGAGAACGAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAANAACC 1020

341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db ACAGATAACGCTTCCACCTTGGATGACAGATTTT 1053

RESULT 8
AB166671 1056 bp mRNA linear PRI 09-MAR-2004
LOCUS Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase
DEFINITION small subunit 2, complete cds.
ACCESSION AB166671
VERSION AB166671.1 GI:45259568
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Ugai,H. and Yokoyama,K.K.
TITLE Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase
JOURNAL Published Only in Database (2004)
REFERENCE 2 (bases 1 to 1056)
AUTHORS Ugai,H. and Yokoyama,K.K.
TITLE Direct Submission
JOURNAL Submitted (05-MAR-2004) Hideyo Ugai, RIKEN BioResource Center, Gene
Engineering Division; 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074,
Japan (E-mail:ugai@rtc.riken.go.jp, Tel:81-29-836-3612,
Fax:81-29-836-9120)
FEATURES
Location/Qualifiers
1..1056
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="p53R2.C19"
/cell_lines="Saos-2"
/cell_type="Osteogenic sarcoma"
1..1056
/gene="p53R2"
1..1056
/gene="p53R2"

gene
CDS

```

```

/codon_start=1
/product="p53-inducible ribonucleotide reductase small
subunit 2"
/protein_id="BAD12267.1"
/db_xref="GI:45259568"
/translation="MGDPPEAAGLQDSSDTNESEIKSNEEP LLKSSRRFVI
FTQYPDIMWKYQQAQASFTABEDLSKLPWNKADKDYITSHLAFPAASDGI
VFNENLVERFQVQVPEARCFYGOIILINSHSEMYSLIDITRDPKRFELFNAIE
TMEYVKKADMLRWIADRKSTGERVVAFAAEGVFPFSGSFAAIFWLKKGLMPLGT
FSNELSRDGLHCDFAFLMFQILVNKPSERVRREIIVDAVKIEQEFTEALPVGLIG
MNCILMKQYIEFVADRLVLLVLFGLSKVFAENPFDFMENISLEGKTNFFEKRVSEYORF
AVMAETTDNVFTLDAF"

ORIGIN
Alignment Scores: 1.01e-170 Length: 1056
Pred. No.: 1821.00 Matches: 351
Score: 1821.00
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-698-228-1 (1-351) x AB166671 (1-1056)
QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 1 ATGGCGACCCGGAAGCCGCGAGCGCGCTGGATCAGATGAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
Db 61 GACACCAACGAAAGTGAATAAAGTCAATCAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Db 121 CGGTTTGTCTATCTTCCATCCAGTACCTCATATTGGAAAAATGTATAAAGCAGCACAG 180
QY 61 AlaSerPheTrpThrAlaGluGluValAspLysSerLysAspLeuProHisTrpAsnLys 80
Db 181 GCTTCTCTTGGACAGCAGAGAGGTGCGACTTATCAAGAGGATCTCCCTCCTCGGAACAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 241 CTTAAGCAGATGAGAAGTACTTCACTCTCAGATCTTAGCTTTTTCAGCCAGTATGAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 301 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGAGGTGCGAGGTTCCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTCTCATCAGATGTACAGTTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTATTAATGCAATTGAA 480
QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCCCTTGGCATGATGATGATGATGATA 540
QY 181 SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
Db 541 TCTACTTTTGGGGAAGAGTGGTGGCCCTTTGCTGCTAGAGAGGAGTCTTCTTCTCAGGA 600
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGTCTGCTATATTCTGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260

```

Db 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTCATGCTGTCAAA 780
 Qy 261 ILeGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 Db 781 ATGTAGCAGGAGTATTAACAGAGCCTTGGCAGTTGGCCTCATTTGGAATGAATTCAT 840
 Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 Db 841 TTGATGAACACGTACATTCAGTTGTAGCTGACAGATTACTTGTGAACTTGGATCTCA 900
 Qy 301 LysValPheGluAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 Db 901 AAGGTTTTTCAGGCAGAAAATCCTTTGATTTTATGGAACAATTTCTTTAGAAAGGAAA 960
 Qy 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 Db 961 ACAAAATTTCTTTGAGAAACAGATTTCAGAGTATCAGCGTTTGCAGTTATGCGAAGAAC 1020
 Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1021 ACAGATAAGCTCTTCACCTTGGATGAGATTTT 1053

RESULT 9

AR454870
 LOCUS AR454870 1081 bp DNA linear PAT 20-FEB-2004
 DEFINITION Sequence 4 from patent US 6682917.
 ACCESSION AR454870
 VERSION AR454870.1 GI:42688825
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1081)

AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
 TITLE Protein having a ribonucleotide Reductase activity and a DNA thereof

JOURNAL Patent: US 6682917-A 4 27-JAN-2004;

FEATURES
 source Location/Qualifiers

1..1081 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Alignment Scores:
 Pred. No.: 1,04e-170 Length: 1081
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-698-228-1 (1-351) x AR454870 (1-1081)

Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
 Db 20 ATGGGGCCCGGAAAGCGGAGCGCGCGCTGGATCAGATGAGATCATCTTCA 79
 Qy 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
 Db 80 GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 139
 Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysGlnAlaGln 60
 Db 140 CGGTTTGTTCATCTTCCATCCAGTACCCTGATATTTGGAAAATGATATAACAGGCACAG 199
 Qy 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisIlePheLys 80
 Db 200 GCTTCCTCTCGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAAACAG 259
 Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAsp 100
 Db 260 CTTAAAGCAGATGAGAAGTACTTCTCATCTTACATCTTAGCCTTTTTCAGCCAGTGAT 319

Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 Db 320 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTCGAGGTCAGAGGCT 379
 Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 Db 380 CGCTGTTTCTATGGCTTCAAATTTCTCATCAGAAATGTTCTCACTCAGAGATGACAGTTTG 439
 Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 Db 440 CTGATAGACACTTACATCAGATATCCAGAAAAGGAAATTTTATTTAATGCAATTGAA 499
 Qy 161 ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
 Db 500 ACCATGCCCTATGTTAAGAAAAGCAGATTGGCCCTTCGATGATAGCAGATAGAAA 559
 Qy 181 SerThrPheGlyGluArgValAlaPheAlaValGluGluValPhePheSerGly 200
 Db 560 TCTACTTTTGGGAAAGAGTGGTGGCCTTTGCTGCTAGAAAGAGTTTCTTCTCAGGA 619
 Qy 201 SerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 Db 620 TCTTTTGTGCTATATTTCTGGCTAAAGAAAGAGAGGCTCTTATGCCAGAGCTCAC 679
 Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db 680 AATGAACCTCATCAGCAGAGATGAGAGCTTCACCTGACCTTTGCTGCTGATGTTCCA 739
 Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 Db 740 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTTGCTGTCAA 799
 Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 Db 800 ATTGACGAGGAGTTTTTAACAGAAAGCCTTGCAGATTTCGCTCATTTGGAATGCAAT 859
 Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 Db 860 TTGATGAACACGTACATTCAGTTGTAGCTGACAGATTACTTGTGGAACCTTGGATCTCA 919
 Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 Db 920 AAGTTTTTTCAGGCAGAAAATCCTTTGATTTTATGAAAACATTTCTTTAGAAAGGAAA 979
 Qy 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 Db 980 ACAAAATTTCTTTCAGAAACAGATTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAAGAAC 1039
 Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1040 ACAGATAAGCTCTTTCACCTTGGATGAGATTTT 1072

RESULT 10

BD064766
 LOCUS BD064766 1081 bp DNA linear PAT 27-AUG-2002
 DEFINITION Novel protein and DNA thereof.
 ACCESSION BD064766
 VERSION BD064766.1 GI:22610369
 KEYWORDS JP 2001269184-A/3.
 SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

1 (bases 1 to 1081)
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.

TITLE Novel protein and DNA thereof

JOURNAL Patent: JP 2001269184-A 3 02-OCT-2001;

YUSUKE NAKAMURA,TAKEDA CHEMICAL INDUSTRIES LTD

OS Homo sapiens (human)

PN JP 2001269184-A/3

PD 02-OCT-2001

PF 27-JUN-2000 JP 2000192401

PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
 C12N15/09, A61K31/711, A61K38/00, A61K38/43, A61K45/00, A61K48/00, PC
 A61P35/00,
 PC A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC
 , C12N5/10, C12N9/02,
 PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53// (C12N15/09,
 PC C12R1:91),
 PC (C12N15/09, C12R1:19), (C12N9/02, C12R1:91), (C12P21/02, C12R1:91),
 PC (C12P21/08, C12R1:91), C12N15/00, A61K37/02, A61K37/48, C12N5/00,
 PC (C12N15/00, C12R1:91), (C12N15/00, C12R1:19)
 CC Novel protein and DNA thereof
 FH Key Location/Qualifiers
 FT source 1..1081
 FT /organism='Homo sapiens (human)'.
 FEATURES
 source
 1..1081
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,04e-170 Length: 1081
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-10-698-228-1 (1-351) x BD064766 (1-1081)
 QY 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
 DB 20 ATGGCGCGACCCGGAAGCGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 79
 QY 21 AspThrAsnGluSerGluLeuLeysSerAsnGluGluProLeuLeuArgLysSerArg 40
 DB 80 GACACCAAGCAAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 139
 QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysGlnAlaGln 60
 DB 140 CGGTTTGTTCATCTTCCCAATCCAGTACCCTGATATTTGGAAATGATATAACAGGCACAG 199
 QY 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
 DB 200 GCTTCCTCTTGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAAACAG 259
 QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisLeuAlaPhePheAlaAlaSerAsp 100
 DB 260 CTTAAAGCAGATGAGAGTACTTCACTCTCACTCTTACCTTTTTCAGCCAGTGAT 319
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 DB 320 GGAATTGTAAATGAAAAATTTGGTGAGCGCTTTAGTCAGAGGTGCGAGGTTCCAGAGGCT 379
 QY 121 ArgCysPheTyrGlyPheGlnIleLeuLeuGluAsnValHisSerGluMetTyrSerLeu 140
 DB 380 CGCTGTTTCTATGGCTTTCAATCTCTCGAGAATGTTCACTCAGAGATGATACAGTTTG 439
 QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 DB 440 CTGATAGACACTTATCATCAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTGAA 499
 QY 161 ThrMetProTyrValLysLysAlaAspTyrPalaLeuArgTyrIleAlaAspArgLys 180
 DB 500 ACCATGCCCTATGTTAAGAAAAAACAGATTCGGCTTCGGATGATACAGATAGAAA 559
 QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 DB 560 TCTACTTTTGGGAAAGAGTGGTGGCTTGTCTGCTAGAGAGGATTTTCTTCTCAGGA 619
 QY 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220

DB 620 TCTTTTGTCTATATTTCTGGCTAAAGACAGAGAGGTCTTATGCCAGGACTCACITTTTCC 679
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 DB 680 AATGAACATCATCAGCAGAGATGAAGGACTTCACCTGTGACTTGTCTGCTGATGTTCCAA 739
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 DB 740 TACTTAGTAATAAGCTTCAGAAAGAGGCTCAGGAGATCATTTGTGCTGTCTCAAA 799
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 DB 800 ATTGAGCAGGAGTTTAAACAGAACCTTCGCCAGTTCGCCTCATTTGGAATGAATGCATT 859
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGlyPheSer 300
 DB 860 TTGATGAAACAGTACATTGAGTTGTAGCTGACAGATTACTTTGGAACTTTGGATTCTCA 919
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 DB 920 AAGTTTTTTCAGCAGCAAAATCCTTTTGATTTTATGAAACATTTCTTTAGAGAGAAA 979
 QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 DB 980 ACAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCCAGAAACC 1039
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1040 ACAGATAACGTCTTTCACCTTGGATGACAGATTTT 1072
 RESULT 11
 BD093079 1081 bp DNA linear PAT 27-AUG-2002
 LOCUS Novel protein and its DNA.
 DEFINITION
 ACCESSION BD093079
 VERSION BD093079.1 GI:22638667
 KEYWORDS WO 0100799-A/3.
 SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1081)
 AUTHORS Nakamura, Y., Arakawa, H. and Tanaka, H.
 TITLE Novel protein and its DNA
 JOURNAL Patent: WO 0100799-A 3 04-JAN-2001;
 TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,
 HIROSHI TANAKA
 COMMENT OS Homo sapiens (human)
 PN WO 0100799-A/3
 PD 04-JAN-2001
 PF 27-JUN-2000 WO 2000JP004189
 PR 28-JUN-1999 JP 99P 181131, 06-JUL-1999 JP 99P 192391 PR
 21-JAN-2000 JP 00P 017770
 PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
 C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC
 , C12P21/02, A61K38/44,
 PC A61K48/00, C07K16/40, G01N33/53, C12Q1/26
 CC
 FH Key Location/Qualifiers
 1..1081
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 FEATURES
 source
 Location/Qualifiers
 Alignment Scores:
 Pred. No.: 1,04e-170 Length: 1081
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

```
US-10-698-228-1 (1-351) x BD093079 (1-1081)
QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
Db 20 ATGGCGCAGCCCGAAGCGCGGCGCTGGATCAGAGTACATCTTCA 79
QY 21 AspThrAsnGluSerGlnLeuLeuSerAsnGluProLeuLeuArgLysSerArg 40
Db 80 GACACCAACGAAAGTAAATGAATGAATGAAGAGCCACTCTTAAGAAAGAGTTCTCCG 139
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysGlnAlaGln 60
Db 140 CGGTTTGTATCTTTCCATCCAGTACCTGATATTTGGAAATGTATATAACAGGCACAG 199
QY 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrPheLys 80
Db 200 GCTTCTCTTGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCAGTGGAAACAAG 259
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 260 CTTAAGCAGATGAGAGTACTTCTCATCTTACATCTTAGCCTTTTTCGACCCAGTGAT 319
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 320 GGAATTTGTAATGAAATTTGGTGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 379
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 380 CGCTGTCTTCTATGGCTTTCAAAATTCATCGAGAATGTTCTCAGAGATGTACAGTTTG 439
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 499
QY 161 ThrMetProTyrValLysLysAlaAspTyrPheAlaLeuArgTyrPheIleAlaAspArgLys 180
Db 500 ACCATGCCCTATGTTAAGAAAGAGAGATTCGGCCTTCGATGAGTATAGATAGAAAA 559
QY 181 SerThrPheGluArgValAlaAlaPheAlaAlaValGluGluValPhePheSerGly 200
Db 560 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTAGAGAGGATTTCTTCTCAGGA 619
QY 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 620 TCTTTTGTCTATATTTCTGGCTAAGAAAGAGAGGTCTTATGCCAGGACTCACTTTTCC 679
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 680 AATGAACCTCATCAGCAGAGATGAAGSACTTCTACTGTGACTTTGCTGCTGATGTTCCAA 739
QY 241 TyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAlaValLys 260
Db 740 TACTTTAGTAAATAAGCCTTCAGAAAGAGGGTCAGGAGATCATGTTGATGCTGTCAAA 799
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 800 ATTGACGAGGATTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAAATGAATTCATT 859
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 860 TTGATGAACAGTACATGAGTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATTTCA 919
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 920 AAGGTTTTTTCGGCAGAAAAATCCTTTGATTTATTTAGAAAAACATTTCTTTAGAAAGAAA 979
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 980 ACAAAATTTCTTTGAGAAACAGGTTTCAGAGTATACAGCGTTTTCAGGATTTATGCAAGAAC 1039
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
|||||
```

```
Db 1040 ACAGATAACGTCTTTCACCTTGGATGCAGATTTT 1072
RESULT 12
AR454869
LOCUS AR454869
DEFINITION Sequence 3 from patent US 6682917.
ACCESSION AR454869
VERSION AR454869.1 GI:42688824
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4955)
AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE Protein having a ribonucleotide Reductase activity and a DNA
thereof
JOURNAL Patent: US 6682917-A 3 27-JAN-2004;
FEATURES Location/Qualifiers
source 1..4955
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores: 7.39e-170 Length: 4955
Pred. No.: 1821.00 Matches: 351
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6
US-10-698-228-1 (1-351) x AR454869 (1-4955)
QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
Db 245 ATGGCGCAGCCCGAAGCGCGGCGCTGGATCAGAGTACATCTTCA 304
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
Db 305 GACACCAACGAAAGTGAATGAATGAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysGlnAlaGln 60
Db 365 CGGTTTGTCTATCTTTCCATCCAGTACCTGATATTTGGAAATGTATAACAGGCACAG 424
QY 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrPheLys 80
Db 425 GCTTCTCTTCTGGACAGAGAGGTGCACTTATCAAGAGATCTCCCTCAGTGGAAACAAG 484
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 485 CTTAAGCAGATGAGAGTACTTCACTCTCACTTACATCTTGGCCTTTTTCGACCCAGTGAT 544
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 545 GGAATTTGTAATGAAATTTGGTGAGCGCTTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 604
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 605 CGCTGTCTTCTATGGCTTTCAAAATTTCTCATCGAATGTTCTCAGATGTACAGATTTG 664
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 724
QY 161 ThrMetProTyrValLysLysAlaAspTyrPheAlaLeuArgTyrPheIleAlaAspArgLys 180
Db 725 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGCCTTGGATGATGATAGATAGAAAA 784
QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db 785 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAAAGAGTCTTCTTCTCAGGA 844
```


QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 Db 845 TCTTTTGTCTGCTATATCTGGCTAAAGAGAGAGAGGCTTTATGCCAGGACTCACTTTTTC 904
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db 905 AATGAACATCATCAGCAGAGATGAAGACATTCATCTGTGACTTTGCTTGCCTCGATGTTCCAA 964
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 Db 965 TACTTAGTAATAAGCTTCAGNAGAAAGGCTCAGGAGATCATTTGTGTGCTGTCAAA 1024
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 Db 1025 ATTGACGAGGAGTTTTTAACAGAGGCTTGGCAGTTGGCTCATTTGGAATGCAATT 1084
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 Db 1085 TTGATGAACAGTACATTCAGTTGTAGCTGACAGATTAATTCTGTGAACTTGGATTCTCA 1144
 QY 301 LysValPheGluAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 Db 1145 AAGTTTTTTCAGCAGAGAAATCTTTTGATTTTATGGAAGAAATTTCTTTAGAGGAGAAA 1204
 QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 Db 1205 ACDAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTTATGGCAGAAACC 1264
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1265 ACAGATAAGCTCTTCACCTTGGATGCAGATTTT 1297

RESULT 13
 BD064765
 LOCUS
 DEFINITION Novel protein and DNA thereof.
 ACCESSION BD064765
 VERSION BD064765.1 GI:22610368
 KEYWORDS JP 2001269184-A/2.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 4955)
 Nakamura,Y., Arakawa,H. and Tanaka,H.
 AUTHORS
 TITLE Novel protein and DNA thereof
 JOURNAL Patent: JP 2001269184-A 2 02-OCT-2001;
 YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD
 COMMENT OS Homo sapiens (human)
 PN JP 2001269184-A/2
 PD 02-OCT-2001
 PF 27-JUN-2000 JP 2000192401
 PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
 C12N15/09, A61K31/711, A61K38/00, A61K38/43, A61K45/00, A61K48/00, PC
 A61P35/00,
 C1 A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC
 , C12N9/10, C12N9/02,
 PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53/// (C12N15/09,
 PC C12R1:91),
 PC (C12N15/09, C12R1:91), (C12N9/02, C12R1:91), (C12P21/02, C12R1:91),
 PC (C12P21/08, C12R1:91), C12N15/00, A61K37/02, A61K37/48, C12N5/00,
 PC (C12N15/00, C12R1:91), (C12N15/00, C12R1:19)
 CC Novel protein and DNA thereof
 FH Key Location/Qualifiers
 FT source 1. .4955
 /organism='Homo sapiens (human)'.
 Location/Qualifiers
 FT
 1. .4955
 /organism='Homo sapiens'
 /mol_type='genomic DNA'
 /db_xref='taxon:9606'

FEATURES
 source
 ORIGIN

Alignment Scores:
 Pred. No.: 7,39e-170 Length: 4955
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-698-228-1 (1-351) x BD064765 (1-4955)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
 Db 245 ATGGCGCACCGGAAAGCGCGGCGGCTGGATCAGGATCAGAGATCATCTTCA 304
 QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
 Db 305 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCCTTAAGAAAGAGTTCTCCG 364
 QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
 Db 365 CGTTTGTCTATCTTTCCATCCAGTACCTGATATTTGGAAATGTATAACAGGCACAG 424
 QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 Db 425 GCTTCCCTTCTGGACAGCAGAGAGGTCGACTTATCAAGAGGATCTCCCTCAGTGAACAAG 484
 QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
 Db 485 CTTAAGCAGATGAGAAAGTACTTCTCATCTCTCACATCTTTAGCCTTTTTCAGCCAGTGAT 544
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 Db 545 GGAATTTGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTCAGGTTCCAGAGGCT 604
 QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 Db 605 CGCTGTTTCTATGCTTCAAAATTTCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 664
 QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTATTAATCAATTTGAA 724
 QY 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 Db 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCGCTTGGCATGATGATAGCAGATAGAAA 784
 QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 Db 785 TCTACTTTTGGGAAAGAGTGGTGGCCTTTGCTGCTGTAGAAGGAGGATTTCTTCTCAGGA 844
 QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 Db 845 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGCTCTTATGCCAGGACTCACTTTTTC 904
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db 905 AATGAACATCATCAGCAGAGATGAAGAGACTTCACCTGTGACTTTGCTTGCCTGATGTTCCAA 964
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 Db 965 TACTTAGTAATAAGCTTCAGNAGAAAGGCTCAGGAGATCATTTGTGTGCTGCTCAAA 1024
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 Db 1025 ATTGACGAGGAGTTTTTAACAGAGGCTTGGCAGTTGGCCTCATTTGGAATGAATTCATT 1084
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 Db 1085 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTAATTCTGTGAACTTGGATTCTCA 1144
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 Db 1145 AAGTTTTTTCAGCAGAGAAATCTTTTGATTTTATGGAACATTTCTTTAGAGGAGAAA 1204


```

QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
|
|
|
Db 1205 ACAAAATTTCTTGGAGAACGATTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1264
|
|
|
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
|
|
|
Db 1265 ACAGATAACGCTTCCACCTTGGATGCGAGATTTT 1297
|
|
|
RESULT 14
BD093078
LOCUS BD093078
DEFINITION Novel protein and its DNA.
ACCESSION BD093078
VERSION BD093078.1 GI:22638666
KEYWORDS WO 0100799-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 4955)
Nakamura, Y., Arakawa, H. and Tanaka, H.
Novel protein and its DNA
TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,
HIROSHI TANAKA
OS Homo sapiens (human)
PN WO 0100799-A/2
PF 27-JUN-2000 WO 2000JP004189
PR 28-JUN-1999 JP 99P 181131, 06-JUL-1999 JP 99P 192391 PR
21-JAN-2000 JP OOP 017770
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC
, C12P21/02, A61K38/44,
CC A61K48/00, C07K16/40, G01N33/53, C12Q1/26
FH Key Location/Qualifiers.
source i. 4955
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 7.39e-170 Length: 4955
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-698-228-1 (1-351) x BD093078 (1-4955)
QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
|
|
|
Db 245 ATGGCGGACCCCGAAGCGCGGAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 304
|
|
|
QY 21 AspThrAsnGluSerGluLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
|
|
|
Db 305 GACACCAAGCAAGTGAATTAAGTCAATGAAGACCACTCTTAAGAAAGAGTTCTCGC 364
|
|
|
QY 41 ArgPheValIlePheProGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
|
|
|
Db 365 CGTTTGTTCATCTTCCATCCAGTACCTGATATTTGGAAATGTAATAACAGGCACAG 424
|
|
|
QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
|
|
|
Db 425 GCTTCTCTTGTGACAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAAACAG 484
|
|
|
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaIleAsp 100
|
|
|

```

```

Db 485 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGTGTAT 544
|
|
|
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
|
|
|
Db 545 GGAATTTGTAATAAGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGAGTTCAGAGGCT 604
|
|
|
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
|
|
|
Db 605 CGCTGTTTCTATGGCTTCAAAATTCATCATCAGAAATGTTCTCATCAGAGATGTACAGTTTG 664
|
|
|
QY 141 LeuIleAspThrTyrIleArgAspProLysArgGluPheLeuPheAsnAlaIleGlu 160
|
|
|
Db 665 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTATTATGCAATTGAA 724
|
|
|
QY 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTTrpIleAlaAspArgLys 180
|
|
|
Db 725 ACCATGCCCTATGTTAAGAAAAAGCAGATGGCGCTTGGATGGATGAGATAGAAAA 784
|
|
|
QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
|
|
|
Db 785 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTAGAGAGGATTTTCTTCTCAGGA 844
|
|
|
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
|
|
|
Db 845 TCTTTTGTCTATATTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 904
|
|
|
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
|
|
|
Db 905 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGCTTGTGCTGCTGATGTTCAA 964
|
|
|
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
|
|
|
Db 965 TACTTAGTAATAAGCTTCAGAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAA 1024
|
|
|
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
|
|
|
Db 1025 ATTGAGCAGGAGTGTAAACAGAAAGCTTGCAGTTCCTCATTTGGAATGAATTCATT 1084
|
|
|
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
|
|
|
Db 1085 TTGATGAACACAGTACATTTAGTTGTAGCTGCACAGATTACTTTGGAACTTTGGATTCTCA 1144
|
|
|
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
|
|
|
Db 1145 AAGTTTTTCAGCAGAAAAATCCTTTGATTTTATGGAATAACATTTCTTTAGAGGAAAA 1204
|
|
|
QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
|
|
|
Db 1205 ACAAAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTTTCAGTTATGCGAGAAACC 1264
|
|
|
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
|
|
|
Db 1265 ACAGATAACGCTTCCACCTTGGATGCGAGATTTT 1297
|
|
|
RESULT 15
BD093063
LOCUS BD093063
DEFINITION Homo sapiens p53R2 mRNA for ribonucleotide reductase, complete cds.
ACCESSION AB036063
VERSION AB036063.1 GI:7229085
KEYWORDS p53R2; ribonucleotide reductase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
Tanaka, H., Arakawa, H., Yamaguchi, T., Shiraishi, K., Fukuda, S.,
Matsui, K., Takei, Y. and Nakamura, Y.
A ribonucleotide reductase gene involved in a p53-dependent
cell-cycle checkpoint for DNA damage
Nature 404 (6773), 42-49 (2000)
JOURNAL 20179179
MEDLINE 10716435
PUBMED

```

REFERENCE 2 (bases 1 to 4955)
 AUTHORS Tanaka,H., Arakawa,H. and Nakamura,Y.
 TITLE Direct Submission
 JOURNAL Submitted (18-DEC-1999) Yusuke Nakamura, University of Tokyo,
 Institute of Medical Science, Human Genome Center, Laboratory of
 Molecular Medicine; 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639,
 Japan [E-mail:yusuke@ims.u-tokyo.ac.jp. Tel:+81-3-5449-5372,
 Fax:+81-3-5449-5433]

FEATURES
 Source
 1..4955
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosomes="8"
 /map="8q23.1"
 1..4955
 /gene="p53R2"
 245..1300
 /gene="p53R2"
 /codon_start=1
 /product="ribonucleotide reductase"
 /protein_id="BA92434.1"
 /db_xref="GI:7229086"
 /translation="MGDPERPEAAGLDQDERSSDTNESEIKSNBEPRLKRSRRFVI
 FPIQYDIIWKYKQQAQFWTAEVDLSKOLPHWNKLKADKRYFISHILAFPAASDGI
 VNENLVERSQEVQPEARCFYGFQILLENVHSEMYSLIDTVIRDPKRELFNAIE
 TMPYVKKADALRWIADRKSTFGERVAFPAAVEGFFSGSFAAIFNLKRGIMPGLT
 FSNELISDEGLHDFACILFQYLVNKPSEERREIIVDAVKIEQFLEALFPVGLIG
 MNCILMKQYIEFVADRLIVELGFSKVQFQENPFDPMENISLEGKTNFFPKRVSQRF
 AVMAETTONVFLDADF"
 4933
 /gene="p53R2"
 /note="22 a nucleotides"

polya_site
 :
 :
 :

ORIGIN
 Alignment Scores:
 Pred. No.: 7,39e-170 Length: 4955
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-10-698-228-1 (1-351) x AB036063 (1-4955)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
 DB 245 ATGGCGACCCGGGAAGCCGGAAGCGCGCGGCTGGATCAGAGATCATCTTCA 304
 QY 21 AspThrAsnGluSerGluLeuIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
 DB 305 GACACCAACGAAAGTGAATAAAGTCAATGAGAGCCACTCCTAAGAAAGAGTTCTCGC 364
 QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
 DB 365 CGTTTGTGATCTTTCCAAATCCAGTACCCTGATATTGGAAATGATATAACAGGCACAG 424
 QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 DB 425 GCTTCTCTCTGCAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGACAAG 484
 QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
 DB 485 CTTAAAGCAGATGAGAAGTACTTCTATCTCATCTTAGCCCTTTTTCAGCCAGTGAT 544
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 DB 545 GGAATTGTAAATGAAAATTGTGTGGAGCGCTTTAGTCAGAGGTCAGGTTCCAGAGGCT 604
 QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 DB 605 CGCTGTTCTATGGCTTTTCAANTTCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 664

QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 DB 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTCAA 724
 QY 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 DB 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCCCTTGCATGCATAGCAGATAGAAA 784
 QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 DB 785 TCTACTTTTGGGAAAAGAGTGGCTTTGCTGCTGTAGAAGGAGTTTTCTTCTCAGGA 844
 QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 DB 845 TCTTTTGTGCTATATTTCTGGCTAAAGAAGAGAGGTCTTATGCCAGGACTCCTTTTTC 904
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 DB 905 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCTGCTGATGTTCCA 964
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 DB 965 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGTCAGGGAGATCATTTGTGATGCTGTCAA 1024
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 DB 1025 ATTGACGACAGAGTTTTTAAACAGAACCCCTGCCAGTTGGCCTCATTTGGAATGAATTGCATT 1084
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGlyPheSer 300
 DB 1085 TTGATGAACAGTACATTTAGTTTGTAGCTGACAGATTACTTTGTGGAACCTTGGATTCTCA 1144
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGlyLys 320
 DB 1145 AAGGTTTTTTCAGGCAGAAAAATCCTTTTGTATGAAAAACAATTTCTTTAGAAGGAAAA 1204
 QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 DB 1205 ACAAAATTTCTTTGAGAAACAGATTTCAGAGTATCAGCGTTTTTGCAAGTTATGCGCAGAAACC 1264
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1265 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1297

Search completed: October 30, 2005, 08:09:50
 Job time : 3484 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 30, 2005, 06:58:12 ; Search time 443 Seconds
(without alignments)
4690.359 Million cell updates/sec

Title: US-10-698-228-1

Perfect score: 1821

Sequence: 1 MGDPERPRAAGLDQDERSSS.....QRFVMAETTDNVFTLDADF 351

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q/cgn2.1/USPTO.spool/US10698228/runat 26102005.100714.5097/app query.fasta 1.519
-DB=N Geneseq 16Dec04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228 @CGN 1.1.352 @runat 26102005.100714.5097 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 16Dec04.*

1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1821	100.0	1053	4	AAF32438	Aaf32438 Human rib
2	1821	100.0	1053	4	AAF32447	Aaf32447 Human rib
3	1821	100.0	1081	4	AAF32440	Aaf32440 Human rib
4	1821	100.0	2596	3	AAAI2411	Aaai2411 CDNA enco
5	1821	100.0	4955	4	AAF32439	Aaf32439 Human rib

6	1821	100.0	4955	13	ADR24210	Adr24210 Breast ca
7	1817	99.8	1601	4	AAH14924	Aah14924 Human cdn
8	1460.5	80.2	1989	3	AAC78111	Aac78111 Human can
9	1460.5	80.2	2215	12	ADK70302	Adk70302 Respirato
10	1460.5	80.2	2216	10	ADJ56536	Adj56536 Murine cd
11	1460.5	80.2	2482	4	AAS44917	Aas44917 Human con
12	1460.5	80.2	2500	6	ABL65414	Ab165414 Lung can
13	1460.5	80.2	2500	6	ABL65517	Ab165517 Lung can
14	1460.5	80.2	2500	6	ABL65859	Ab165859 Lung can
15	1460.5	80.2	2500	8	ABX10335	Abx10335 DNA encod
16	1460.5	80.2	2500	11	ADI32044	Adi32044 Human cdn
17	1460.5	80.2	2500	12	ADNO4443	Adn04443 Antipsori
18	1460.5	80.2	2500	12	ADQ09273	Adq09273 Human RRM
19	1460.5	80.2	2500	13	ACN37637	Acn37637 Tumour-as
20	1460.5	80.2	3294	4	AAH73225	Aah73225 Human cer
21	1460.5	80.2	3393	12	ADN03788	Adn03788 Antipsori
22	1460.5	80.2	3393	12	ADO19225	Ado19225 Human PRO
23	1447.5	79.5	2641	4	AAS44745	Aas44745 Human ful
24	1429	78.5	1170	2	AAV05641	Aav05641 Human rib
25	1425.5	78.3	1328	12	ADO57308	Ado57308 DNA encod
26	1376.5	75.6	1371	5	AAS79474	Aas79474 DNA encod
27	1348	74.0	1289	4	ABL14627	Ab114627 Drosophil
28	1241.5	68.2	3945	4	ABL14626	Ab114626 Drosophil
29	1230	67.5	1218	8	ABT20895	Abt20895 Aspergill
30	1220.5	67.0	1292	13	ADS49379	Adg49379 Bacterial
31	1210	66.4	14176	2	AAH84564	Aat84564 Swinepox
32	1210	66.4	14176	4	AAF84949	Aaf84949 Reverse c
33	1210	66.4	14176	4	AAF84948	Aaf84948 Nucleotid
34	1192.5	65.5	1146	8	ABT19075	Abt19075 Aspergill
35	1192	65.5	1242	6	ABZ32250	Abz32250 Candida a
36	1188	65.2	706	13	ADQ57092	Adq57092 Novel can
37	1176.5	64.6	963	13	ADS58378	Adg58378 Bacterial
38	1168.5	64.2	1200	10	ACC61238	Acc61238 Gene sequ
39	1168.5	64.2	1200	10	ADK63241	Adk63241 Disease t
40	1168.5	64.2	1200	13	ADT47419	Adt47419 Bacterial
41	1161.5	63.8	1173	13	ADT47814	Adt47814 Bacterial
42	1161.5	63.8	1206	10	ADB69740	Adb69740 C. neofo
43	1151	63.2	1314	8	ABT18481	Abt18481 Aspergill
44	1151	63.2	1450	8	ABT20297	Abt20297 Aspergill
45	1151	63.2	3314	8	ABT17887	Abt17887 Aspergill

ALIGNMENTS

RESULT 1

AAF32438
ID AAF32438 standard; cdna; 1053 BP.

XX AAF32438;

AC AAF32438;

XX 18-APR-2001 (first entry)

DT Human ribonucleotide reductase TP53R2H nucleotide sequence SEQ ID NO:2.

DE Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.

XX Homo sapiens.

XX WO200100799-A1.

PD 04-JAN-2001.

XX 27-JUN-2000; 2000WO-JP004189.

XX 28-JUN-1999; 99JP-00181131.

PR 05-JUL-1999; 99JP-00192391.

PR 21-JAN-2000; 2000JP-00017770.

XX (TAKE) TAKEDA CHEM IND LTD.

XX (NAKA/) NAKAMURA Y.

XX Nakamura Y, Arakawa H, Tanaka H;

DR WPI; 2001-112446/12.
 XX P-PSDB; AAB69050.

PT Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
 PT diagnosis, treatment and prevention of cancer.

XX Claim 5; Fig 1-3; 102pp; Japanese.

PS The present sequence encodes a human ribonucleotide reductase designated
 CC TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of the DNA
 CC repair mechanism and its activity is induced by p53. It can be used for
 CC the treatment, prevention and diagnosis of a wide range of cancers

XX SQ Sequence 1053 BP; 314 A; 192 C; 246 G; 301 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.17e-212 Length: 1053
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-10-698-228-1 (1-351) x AAF32438 (1-1053)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
 DB 1 ATGGGGGACCCGGGAAGCGCGGAGCGCGCGCTGGATCAGATGAGATCATCTTCA 60
 QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
 DB 61 GACACCAACGAAAGTCAATTAAGTCAATGAAGAGCCACTCTCAAGAAAGAGTCTCGC 120
 QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
 DB 121 CGGTTTGTCTATCTTCCCAATCCAGTACCTCCCTGATATTTGGAAAATGATATAACAGGCACAG 180
 QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 DB 181 GCTTCCTCTGGACACAGAGAGGTCGACTTCAAGAGTCTCCCTCACTGGAACAAG 240
 QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAlaSerAsp 100
 DB 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTTACCTTTTTCAGCCAGTGAT 300
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 DB 301 GGAAATGTAAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
 QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 DB 361 CGCTGTTTCTATGGCTTCAAAATTCATCGAGAAGTGTTCACATCAGAGATGTACAGTTG 420
 QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 DB 421 CTGATAGACATACATCAGATCCAGAAAGGGAATTTTATTTATGCAATGAA 480
 QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 DB 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCGCTTGGATGGATAGCAGATAGAAA 540
 QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 DB 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGGCTGTAGAGAGAGTCTTCTCTCAGGA 600
 QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 DB 601 TCTTTTGTCTATATCTGGCTTAAGAGAGAGAGGTCCTATGCGAGACTCACTTTTTC 660
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 DB 661 AATGAACATCATCAGCAGATGAAGACTTCACTGTGACTTTGTGCTTGCCTGATGTTCCAA 720

QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 DB 721 TACTTAGTAATAAGCCTTCAGAGAAGAGGTCAGGGAGATCATTTGTTGATGCTGTCAA 780
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 DB 781 ATTGAGCAGGAGTATTTAAACAGAAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTGCATT 840
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 DB 841 TTGATGAAAACAGTACATTTGAGTTTGTAGCTCAGAGATTACTTGTGGAACCTTGGATTTCA 900
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 DB 901 AAGGTTTTTCAGGCAGAAATCTTTTGATTTTATGGAAACATTTCTTTAGAGAGAAA 960
 QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 DB 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1020
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1021 ACAGATAACGCTCTTCACCTTGGATGCGAGATTTT 1053

RESULT 2

ID AAF32447 standard; cDNA; 1053 BP.
 XX AAF32447;

AC AAF32447;

DT 18-APR-2001 (first entry)

DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:12.
 XX Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.

OS Homo sapiens.

XX WO200100799-A1.

PN 04-JAN-2001.

PD 27-JUN-2000; 2000WO-JP004189.

PF 28-JUN-1999; 99JP-00181131.

PR 06-JUL-1999; 99JP-00192391.

PR 21-JAN-2000; 2000JP-00017770.

XX (TAKE) TAKEDA CHEM IND LTD.

PA (NAKA/) NAKAMURA Y.

PA Nakamura Y, Arakawa H, Tanaka H;

PI WPI; 2001-112446/12.

XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for

PT diagnosis, treatment and prevention of cancer.

XX Claim 6; Page 96-97; 102pp; Japanese.

XX The present invention describes a human ribonucleotide reductase

CC designated TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of

CC the DNA repair mechanism and its activity is induced by p53. It can be

CC used for the treatment, prevention and diagnosis of a wide range of

CC cancers. The present sequence represents a human ribonucleotide reductase

CC related sequence which is given in the exemplification of the present

CC invention

XX SQ Sequence 1053 BP; 314 A; 191 C; 246 G; 302 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.17e-212 Length: 1053

Score: 1821.00 Matches: 351

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0
US-10-698-228-1 (1-351) x AAF32447 (1-1053)			
Qy	1	MetGlyAspProGluuArgProGluAlaLaAGlyLeuAspGlnAspGluArgSerSer	20
Db	1	ATGGCGCAGCCCGAAGGCGGAGCGCGCGCTGGATCAGGATCAGAGATCATCTTCA	60
Qy	21	AspThrAsnGluSerGluuIleLysSerAsnGluGluProLeuLeuArgLysSerArg	40
Db	61	GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAGAGTTCTCGC	120
Qy	41	ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln	60
Db	121	CGGTTTGTTCATCTTCCATCCAGTACCTGCTGATATTTGGAAATGTATTAACAGGCACAG	180
Qy	61	AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys	80
Db	181	GCCTCCTTCGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAAG	240
Qy	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePhealalaSerAsp	100
Db	241	CTTAAAGCAGATGGAAGTACTTTCATCTCACTCTTACGCTTTTTCGAGCCAGTGAT	300
Qy	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluuAla	120
Db	301	GGATTTGTAAATGAANAATTTGTGTGAGGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGT	360
Qy	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140
Db	361	CGCTGTTTCTATGGCTTTTCAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG	420
Qy	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160
Db	421	CTGATAGACACTTATCATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATGAA	480
Qy	161	ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys	180
Db	481	ACCATGCCCTATGTTTAAAGAAAAGCAGATTTGGGCTTTCGATGGATAGCAGATAGAAA	540
Qy	181	SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly	200
Db	541	TCTACTTTTGGGAAAGAGTGTGTGCTTGTGTGTAGAAGGAGTTTCTTCTCAGGA	600
Qy	201	SerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer	220
Db	601	TCCTTTTGTGCTATATTCTGGCTAAAGAAGAGAGTCTTATGTCAGGAGCTCACTTTTTC	660
Qy	221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln	240
Db	661	AATGAACCTCATCAGCAGAGATGAAGCATCTCACTGTGACTTTGCTGCTGATGTTCCAA	720
Qy	241	TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys	260
Db	721	TACTTAGTAATAGCCTTTCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA	780
Qy	261	IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle	280
Db	781	ATTGAGCAGGAGTTTTAAACAGAGCCTTGCAGTGGCTCATTTGGAATGAATGCAAT	840
Qy	281	LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer	300
Db	841	TTGATGAAACAGTACATTTGAGTTTGTAGCTGACATTAATTTGTGGAATCTTGATTTCTCA	900
Qy	301	LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys	320
Db	901	AAGGTTTTTCAGGCAGAAAATTCCTTTGATTTTATGGAANAACATTTCTTTAGNAGGAAA	960
Qy	321	ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr	340

Db	961	ACAAATTCCTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCGAGTTATGGCAGAAACC	102
Qy	341	ThrAspAsnValPheThrLeuAspAlaAspPhe	351
Db	1021	ACAGATAACGTCCTTCACCTTGGATGCAGATTTT	1053
RESULT 3			
AAF32440			
ID	AAF32440	standard; cDNA; 1081 BP.	
XX	AC	AAF32440;	
XX	XX		
XX	DT	18-APR-2001 (first entry)	
XX	DE	Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:4.	
XX	XX	Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.	
KW	XX	Homo sapiens.	
OS	XX	WO200100799-A1.	
XX	XX	04-JAN-2001.	
PD	XX	27-JUN-2000; 2000WO-JP004189.	
XX	XX	28-JUN-1999; 99JP-00181131.	
PR	XX	06-JUL-1999; 99JP-00192391.	
PR	XX	21-JAN-2000; 2000JP-00017770.	
XX	XX	(TAKE) TAKEDA CHEM IND LTD.	
PA	XX	(NAKA/) NAKAMURA Y.	
XX	XX	Nakamura Y, Arakawa H, Tanaka H;	
PI	XX	WPI; 2001-112446/12.	
XX	XX	Ribonucleotide reductase involved in DNA repair and DNA encoding it, for	
PT	XX	diagnosis, treatment and prevention of cancer.	
PS	XX	Example 2; Page 91; 102pp; Japanese.	
XX	XX	The present invention describes a human ribonucleotide reductase	
CC	CC	designated TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of	
CC	CC	the DNA repair mechanism and its activity is induced by p53. It can be	
CC	CC	used for the treatment, prevention and diagnosis of a wide range of	
CC	CC	cancers. The present sequence represents a human ribonucleotide reductase	
CC	CC	related sequence which is used in an example from the present invention	
XX	XX	Sequence 1081 BP; 322 A; 199 C; 252 G; 308 T; 0 U; 0 Other	
SQ			

Alignment Scores:					
Pred. NO.:	4.34e-212	Length:	1081		
Score:	1821.00	Matches:	351		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	4	Gaps:	0		
US-10-698-228-1 (1-351) x AAF32440 (1-1081)					
Qy	1	MetGlyAspProGluArgProGluAlaLaGlyLeuAspGlnAspGluArgSerSer	20		
Dd	20	ATGGCGCAGCCGGAAAGCCGGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA	79		
Qy	21	AspThrAenGluSerGluIleLysSerAenGluGluProLeuLeuArgLysSerSerArg	40		
Dd	80	GACACCACGAAGAAGTAANTAAAGTCANTAGAGGCCACTCTCTAAGAAGAAGTTCTCGC	139		
Qy	41	ArgPheValIlePheProIleGlnTyrrProAspllleTrpLysMetTyrllysGlnlaaGln	60		
Dd	140	CGGTTTGTCATCTTTCCAAATCCAGTACCCTGATATTGGAAAATGTATAAACAGGCACAG	199		

QY 61 AlaSerPheThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 DB 200 GCTTCCTCTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCTCTCACTGGACAAG 259
 QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
 DB 260 CTTAAAGCAGATGAGAGTACTCTCTCATCTTACATCTTACGCTTTTTCGAGCCAGTGAT 319
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 DB 320 GGAATGTAAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGGTTTCAGAGGCT 379
 QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 DB 380 CGCTGTTCTATGGCTTTCAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 439
 QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 DB 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 499
 QY 161 ThrMetProTyrValLysLysLysAlaAspTTPAlaLeuArgTTPAlaAspArgLys 180
 DB 500 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCCCTTCGATGGATAGACATAGAAA 559
 QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 DB 560 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCAGGA 619
 QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 DB 620 TCTTTTGTCTATATCTCTGGCTAAAGAGAGAGGCTTTATGCGCAGCATCACTTTTCC 679
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 DB 680 AATGAACCTATCAGCAGAGATCAGAGACTTCTACTGTGACTTTGCTTGCCTGATGTTCCAA 739
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 DB 740 TACTTAGTAAATAAGCTTTCAGAGNAAGGCTCAGGAGATCATTTGTTGATGCTGCMAA 799
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 DB 800 ATTGACGAGGAGTTTAAACAGAGCCTTGCCAGTGTGCTCATTCGATGAATGCAATT 859
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 DB 860 TTGATGAACACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGGAATTTGGATTTCTCA 919
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 DB 920 AAGGTTTTTCAGCGAGAAATCTTTTGATTTATGGAACAATTTCTTTAGAGGAA 979
 QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 DB 980 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGCTTATGGCAGAAACC 1039
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1040 ACAGATAAGCTTTCACCTTGGATGCAGATTTT 1072

RESULT 4
 AAA12411
 ID AAA12411 standard; cDNA; 2596 BP.

XX AAA12411;
 XX 25-JUL-2000 (first entry)
 XX cDNA encoding a human RNA-associated protein.
 XX Human; RNA-associated protein; cell proliferation; cancer; inflammation;
 KW immune response; reproductive disorder; actinic keratosis;
 KW atherosclerosis; arteriosclerosis; bursitis; cirrhosis; hepatitis;

KW mixed connective tissue disease; myelofibrosis; primary thrombocythemia;
 KW paroxysmal nocturnal hemoglobinuria; polycythemia vera; psoriasis;
 KW trauma; ss.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT 56..1111
 FT CDS /*tag= a
 FT /*product= "RNA-associated protein"

XX WO200015799-A2.

PN 23-MAR-2000.

XX 17-SEP-1999; 99WO-US021688.

XX 17-SEP-1998; 98US-00156039.

PR 22-SEP-1998; 98US-00158720.

PR 04-NOV-1998; 98US-00186815.

PR 08-APR-1999; 99US-0128660P.

XX (INCY-) INCYTE PHARM INC.

XX Tang YT, Corley NC, Guegler KJ, Gorgone GA, Patterson C;
 PI Hillman JL, Baughn MR, Lal P, Azimzai Y, Yue H, Yang J;

XX WPI; 2000-271437/23.

DR P-PSDB; AAY84439.

XX New peptides and polynucleotides, useful for preventing and treating
 PT a disorder associated with increased or decreased expression of RNA
 PT associated proteins.

XX Claim 9; Page 120-121; 131pp; English.

XX The present sequence encodes a human RNA-associated protein. The
 CC expression of RNA-associated proteins is closely associated with
 CC reproductive tissues, nervous tissues, cell proliferation including
 CC cancer, inflammation and immune responses, and so they may be used for
 CC diagnosis, treatment or prevention of cell proliferative,
 CC immune/inflammatory disorders, and reproductive disorders. Diseases and
 CC disorders which may be treated include actinic keratosis,
 CC atherosclerosis, arteriosclerosis, bursitis, cirrhosis, hepatitis, mixed
 CC connective tissue disease, myelofibrosis, paroxysmal nocturnal
 CC hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia
 CC and cancers, and trauma

XX SQ Sequence 2596 BP; 783 A; 418 C; 510 G; 885 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.59e-211	Length:	2596
Score:	1821.00	Matches:	351
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	3	Gaps:	0

US-10-698-228-1 (1-351) x AAA12411 (1-2596)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20

DB 56 ATGGCGACCCCGAAGAGCGCGGCTGGATCAGATGAGAGATCACTTCA 115

QY 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40

DB 116 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCG 175

QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60

DB 176 CGGTTTGTCTATCTTTCCATCCAGTACCTCATATTGGAAATGTATAAACAGGCACAG 235

QY 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80

```
Db 236 GCTTCTCTTGACAGCAGAGAGGTGCACTTATCAAGGATCTCCTCACTGGAACAAG 295
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaIleSerAsp 100
Db 296 CTTAAAGCAGATGAGAAGTACTTCACTCTCATCTTACATCTTACCCCTTTTTCGAGCCAGTGAT 355
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 356 GGAATGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 415
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 416 CGCTGTTCTATGGCTTTCAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 475
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 476 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 535
Qy 161 ThrMetProTyrValLysLysValAlaAspTTPAlaLeuArgTyrPheAlaAspArgLys 180
Db 536 ACATGCCCTATGTTAAGAAAGAGAGATGGGCTTCGATGGATGAGATAGAGAA 595
Qy 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db 596 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGA 655
Qy 201 SerPheAlaAlaIlePheTyrPheLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 656 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 715
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 716 AATGAACATCAGCAGAGATGAAGACATCACTGTGACTTTGCTTGCCTGATGTTCCAA 775
Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 776 TACTTAGTAAATAGCTTCAGAGNAGGGTTCAGGAGATCATTTGTTGATGCTGTCAA 835
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 836 ATTGAGCAGAGTTTTAAACAGAGCCTTGCAGTTGGCCTCATTGGAATGAATTCATT 895
Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 896 TTGATGAAACACATACATTGAGTTTGTAGCTGACAGATTACTTTGTGGAACCTTGGATTCTCA 955
Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 956 AAGGTTTTTCAGCAGAGAAATCCTTTTGATTTATGGAACAATTTCTTTAGAAGGAAA 1015
Qy 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 1016 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGATTATGCAAGAAC 1075
Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1076 ACAGATAACGTCTTCACCTTGGATGAGATTTT 1108
```

RESULT 5

AAF32439
ID AAF32439 standard; cDNA; 4955 BP.

XX AC
XX AAF32439;

XX DT 18-APR-2001 (first entry)

XX DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:3.

XX KW Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.

XX OS Homo sapiens.

PN WO200100799-A1.
XX 04-JAN-2001.
XX 27-JUN-2000; 2000WO-JP004189.
XX 28-JUN-1999; 99JP-00181131.
PR 06-JUL-1999; 99JP-00192391.
XX 21-JAN-2000; 2000JP-00017770.
XX (TAKEDA) TAKEDA CHEM IND LTD.
PA (NAXA/) NAKAMURA Y.
XX Nakamura Y, Arakawa H, Tanaka H;
XX WPI; 2001-112446/12.
DR PT Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
PT diagnosis, treatment and prevention of cancer.
XX Example 2; Page 87-90; 102pp; Japanese.

CC The present invention describes a human ribonucleotide reductase
CC designated TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of
CC the DNA repair mechanism and its activity is induced by p53. It can be
CC used for the treatment, prevention and diagnosis of a wide range of
CC cancers. The present sequence represents a human ribonucleotide reductase
CC related sequence which is used in an example from the present invention
XX

SQ Sequence 4955 BP; 1526 A; 803 C; 999 G; 1627 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4,13e-211 Length: 4955
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-698-228-1 (1-351) x AAF32439 (1-4955)

Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
Db 245 ATGGGCGACCCGGAAAGCCGGAAGCCGGCGGTGGATCAGATGAGAGATCATCTTCA 304
Qy 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
Db 305 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 364
Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
Db 365 CGGTTTGTTCATCTTCCAAATCCAGTACCTGATATTTGGAAAAATGTATAACAGGCACAG 424
Qy 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
Db 425 GCTTCTCTTGGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 484
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 485 CTTAAAGCAGATGAGAAGTACTTCACTCTCATCTTACATCTTGTTCGAGCCAGTGAT 544
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 604
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 605 CGCTGTTTCTATGGCTTTCAAATTCATCAGAAATGTTCTCAGAGATGTTACAGATTG 664
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 724

QY 161 ThrMetProTyrValLysLysLysAlaAspTTPalaLeuArgTrpIleAlaAspArgLys 180
 Db |||||
 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCCCTTCGGATGGATAGACAGATAGAAA 784
 QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 Db |||||
 785 TCTACTTTTGGGGAAGAGCTGGTGGCTTTGCTGCTGTAGAAAGAGTTCCTTCAGGA 844
 QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 Db |||||
 845 TCTTTTGGCTATATCTGGCTAAAGAGAGAGGCTTATGCCAGAGCTCACTTTTCC 904
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db |||||
 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 964
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAlaValLys 260
 Db |||||
 965 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGTCAGGGAGATCATTTGTGCTGTCAA 1024
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 Db |||||
 1025 ATTGACGAGAGTTTTTACAGAGGCTTCGCGAGTTGGCTCATTTGGAATGAATTCGATT 1084
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 Db |||||
 1085 TTGATGAACAGTACATTCAGTTGTAGCTGACAGATTACTTCTGGAACCTTGGATTCCTCA 1144
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 Db |||||
 1145 AAGGTTTTTTCAGGCAGAAAATCCTTTTGATTTTATGGAACAACATTTCTTTAGAAAGGAAA 1204
 QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 Db |||||
 1205 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGATTATGGCAGAAACC 1264
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db |||||
 1265 ACAGATAAGCTCTTACCTTGGATGCAGATTTT 1297

RESULT 6

ID ADR24210 standard; DNA; 4955 BP.

AC ADR24210;

DT 21-OCT-2004 (first entry)

XX Breast cancer prognosis marker #71.

DE ds; breast cancer; prognosis; gene expression; diagnosis.

KW Homo sapiens.

OS WO2004065545-A2.

PN 05-AUG-2004.

XX 15-JAN-2004; 2004WO-US001100.

PR 15-JAN-2003; 2003US-00342887.

XX (ROSE-) ROSETTA INPHARMATICS LLC.

PA (NECA-) NETHERLANDS CANCER INST.

XX Van't Veer LJ, He Y;

XX WPI; 2004-593473/57.

XX Classifying a breast cancer patient according to prognosis comprises
 PT determining the similarity between the level of expression of each of
 PT five genes in a cell sample taken from patient, to control levels.

PS Disclosure; SEQ ID NO 71; 226pp; English.
 CC
 CC The invention relates to a method of classifying a breast cancer patient
 CC according to prognosis by determining the similarity between the level of
 CC expression of each of five genes for which markers are listed in the
 CC specification, in a cell sample taken from the breast cancer patient, to
 CC control levels of expression for each respective five genes to obtain a
 CC patient similarity value. The methods are useful for classifying a breast
 CC cancer patient according to prognosis. Kits and computer program products
 CC are useful for data analysis using the diagnostic, prognostic and
 CC statistical methods of the invention. This sequence corresponds to a
 CC marker used in the method of the invention.
 XX

SQ Sequence 4955 BP; 1526 A; 803 C; 999 G; 1627 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4,13e-211 Length: 4955
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 13 Gaps: 0

US-10-698-228-1 (1-351) x ADR24210 (1-4955)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
 Db 245 ATGGCGCAGCCGGAAGCCGGAAGCGCGCGCTGGATCAGATGAGAGATCATCTTCA 304
 QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
 Db 305 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTCTCGC 364
 QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
 Db 365 CGGTTTGTCTACTTTCATCCAGTACCTGATATTTGGAAAATGTATAAACAGGCACAG 424
 QY 61 AlaSerPheTrpThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 Db 425 GCTTCCTCTTGGCAGCAGAAAGAGTTCGACTTATCAAGAGATCTCCCTCATCTGGAACAAG 484
 QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAspAsp 100
 Db 485 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT 544
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 Db 545 GGAATTGTAAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGTCAGAGGCT 604
 QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 Db 605 CGCTGTTTCTATGGCTTTCAAAATTCATCAGAAATGTTTCTACTCAGAGATGATACAGTTTG 664
 QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 Db 665 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTCAA 724
 QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 Db 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCTTGGCATGCATGATAGATAGAAAA 784
 QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 Db 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGATTTCTTCTCAGA 844
 QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 Db 845 TCTTTTGGCTGCTATATTTCTGGCTAAAGAGAGGCTCTTATGCCAGAGCTCACTTTTCC 904
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGCTGATGTTCCAA 964

Qy	241	TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys	260
Db	965	TACTTAGTAAATTAAGCCCTTCAGAAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAA	1024
Qy	261	IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle	280
Db	1025	ATTGAGCAGGAGTTTAAACAGAAGCCCTGCCAGTTGGCCTCAATTGGAAATGAATTGCATT	1084
Qy	281	LeuMetLysGlnTyrlleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer	300
Db	1085	TTGATGAACAGATACATTGAGTTTGAGCTGACAGATATTCTGTGGAACTTGGATTCTCA	1144
Qy	301	LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys	320
Db	1145	AAGTTTTTCAGGCAGAAATCTTTTGATTTTATGGAACAATCTTTTAGAAGGAAAA	1204
Qy	321	ThrAsnPhePheGluLysArgValSerGluTyxGlnArgPheAlaValMetAlaGluThr	340
Db	1205	ACAAATTTCTTTCAGAAACCGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC	1264
Qy	341	ThrAspAsnValPheThrLeuAspAlaAspPhe	351
Db	1265	ACAGATAACGCTTTCACCTTGGATGCAGATTTT	1297

RESULT 7

AAH14924
ID AAH14924 standard; cDNA; 1501 BP.
XX
XX
AAH14924;
XX AC
XX
26-JUN-2001 (first entry)
XX DT
XX
Human cDNA sequence SEQ ID NO:12810.
DE
XX
Human; primer; detection; diagnosis; antisense therapy; gene therapy; sa.
KW

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

PS Claim 8; SEQ ID NO 12810; 2537pp + Sequence Listing; English.

The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a

QY 218 rPheSerAenGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 Db 845 ATTTTCTAATGAATATTATAGCAGAGATGAGGGTTTACATGTTGATTTTGGCTTGCCTGAT 904
 QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAl 258
 Db 905 GTTCAACACCTGGTACCAACCACTCGGAGGAGAGAGTAAGAGAATAATATATCAATGC 964
 QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 Db 965 TGTTCGATAGAACAGGAGTTCTACTAGGCGCTTGGCTTGAACCTCATTTGGGATGAA 1024
 QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG1 298
 Db 1025 TTGCACCTCTAATGAAGCAATACATTAGATTGTGGCAGACAGACTTATGCTGGAATGGG 1084
 QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
 Db 1085 TTTTAGCAAGGTTTTCAGAGTAGAGAACCACTTGTACTTTATGGAGATATTTCACTGGA 1144
 QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetal 338
 Db 1145 AGGAAGACTAATCTTTTGAAGAGAGATGCGCGAGTATCAGAGATCGGAGTGATGTC 1204
 QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1205 AGTCCACACAGAGATTTCTTTACCTTGGATGCTGACTTC 1244

RESULT 9

ADK70302
 ID ADK70302 standard; cDNA; 2215 BP.
 AC ADK70302;
 XX
 DT 06-MAY-2004 (first entry)
 DE Respiratory disease differentially expressed cDNA #38.
 XX
 KW ds; gene; cytostatic; respiratory; antiasthmatic; gene therapy;
 KW differential gene expression; respiratory disorder; lung cancer;
 KW chronic obstructive pulmonary disease; emphysema; asthma.
 XX
 OS Homo sapiens.
 XX
 PN WO2003101283-A2.
 XX
 PD 11-DEC-2003.
 XX
 PF 02-JUN-2003; 2003WO-US017409.
 XX
 PR 04-JUN-2002; 2002US-0386003P.
 XX
 PA (INCY-) INCYTE CORP.
 XX
 PI Rickert PK, Krasnow R;
 XX
 DR WPI; 2004-042945/04.
 XX
 PT New combination comprising cDNAs and proteins that are differentially
 PT expressed in respiratory disorders, useful for diagnosing or treating
 PT respiratory diseases e.g. lung cancer, chronic obstructive pulmonary
 PT diseases or asthma.
 XX
 PS Claim 1; SEQ ID NO 38; 343pp; English.
 XX
 CC The invention relates to cDNA sequences that are differentially expressed
 CC in respiratory disorders or their complements or encoded proteins. The
 CC cDNAs and proteins are useful for diagnosing, treating or monitoring
 CC treatment of a subject with a respiratory disease including lung cancer,
 CC chronic obstructive pulmonary diseases, emphysema or asthma. The protein
 CC is also useful for screening molecules or compounds to identify at least
 CC one ligand which specifically binds the protein. It is also useful for
 CC preparing and purifying a polyclonal or monoclonal antibody. This

CC sequence corresponds to a cDNA of the invention.
 XX
 SQ Sequence 2215 BP; 593 A; 476 C; 523 G; 623 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,42e-167 Length: 2215
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 12 Gaps: 1

US-10-698-228-1 (1-351) x ADK70302 (1-2215)

QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
 Db 188 CCTGAGCGGACCGCGCTCTGGCCAGCAAGACCGCGAGGAGGATCTTCCAGGAGCCAC 247
 QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
 Db 248 GGAGCCGAAACCTTAAGCAGCTGCCCGCGGTGGAGATGAGCGCTGCTGAGAGAAA 307
 QY 38 rSerArgArgPheValIlePhePheProIleGlnTyrProAspIleTyrLysMetTyrLysG1 58
 Db 308 CCCCCCGCGCTTGTTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGATATAAGAA 367
 QY 58 nAlaGlnAlaSerPheThrThrAlaGluGluValAspLeuSerLysAspLeuProHigTr 78
 Db 368 GGCAGAGGCTTCTCTTTGGACCGCGAGGAGGTGGACCTCTCCAAGGACATTCAGCAGCTG 427
 QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
 Db 428 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCCATGCTTCTGGCTTCTTTTGGAGC 487
 QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
 Db 488 AAGCGATGGCATAGTAATGAAACCTTGTGGAGCGGATTAGCCAGAGAGTTTCAGATTAC 547
 QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
 Db 548 AGAAGCCCGCTGTTCTATGGCTTCCAAATTTGCCATGCAAAACATACATCTCTGAAATGTA 607
 QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
 Db 608 TAGTCTTCTTATTGACACTTACATAAAAGATCCCAAGAAAGGGAATTTCTTCTCAATGC 667
 QY 158 aIleGluThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrIleAlaAs 178
 Db 668 CATTTGAACGATGCTTGTGTCAAGAGAGAGGCGAGACTGGGCTTGGCTGGATTGGGA 727
 QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePh 198
 Db 728 CAAAGAGGCTACCTATGTTGAACTGTTGTAGCTTGTGCTGAGTGAAGGCAATTTCTT 787
 QY 198 eSerGlySerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 Db 788 TTCCGGTCTTTTGGCTCGATATTTCTGGCTCAAGAAACGAGGACTGATGCTCGGCTCAC 847
 QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 Db 848 ATTTTCTAATGAACCTTATACAGAGATGAGGGTTTACATGATGATTTTGTGCTGCTGAT 907
 QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
 Db 908 GTTCAACACCTCGTACACAAACCATCGGAGGAGAGAGTAAGAGAAATAATATCAATGC 967
 QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 Db 968 TGTTCGATAGAACAGGAGTTCTCTAGGCGCTTCCCTGTGAGAGCTCATTTGGGATGAA 1027
 QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG1 298
 Db 1028 TTGCACCTCTAATGAAGCAATATATGATTTTGGCAGACAGACTTATGCTGGAACTGGG 1087

QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
 DB 1088 TTTTACAGAGGTTTTCAGAGTAGAGAACCCATTTGACTTATGGAGAAATTTCACTGGA 1147
 QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
 DB 1148 AGGAAAGACTAACTCTTTTGAGAGAGAGTAGCGAGTATCAGAGGATGGAGTGATGTC 1207
 QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1208 AAGTCAACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1247

RESULT 10

ID ADJ56536
 XX ADJ56536 standard; cDNA; 2216 BP.

AC ADJ56536;

XX 06-MAY-2004 (first entry)

DE Murine cDNA differentially expressed in MYCN activated cells SeqID 342.

XX mouse; murine; differential expression; transactivator; proto-oncogene;
 KW neuroblastoma; small cell lung cancer; cytostatic; gene therapy; ss;
 KW MYCN activated cell.

XX Mus musculus.

OS US2003119009-A1.

XX 26-JUN-2003.

XX 25-FEB-2002; 2002US-00084817.

XX 23-FEB-2001; 2001US-0270784P.

XX (STUA/) STUART S G.

XX (NUCH/) NUCHTERN J G.

XX (PLON/) PLON S E.

XX (SHOH/) SHOHET J M.

XX Stuart SG, Nuchtern JG, Plon SE, Shohet JW;

XX WPI; 2003-635698/60.

XX New genes regulated by MYCN activation, useful in gene therapy,
 PT particularly for treating a subject with e.g. neuroblastoma or
 PT cancers, or for diagnosing, staging or monitoring the treatment of the
 PT cancer.

XX Claim 1; SEQ ID NO 342; 27pp; English.

XX This invention relates to novel isolated cDNAs that are differentially
 CC expressed in MYCN activated cells. Specifically, it refers to
 CC polynucleotide sequences that exhibit differential expression patterns in
 CC cells activated by the transactivator MYCN, where MYCN is a proto-
 CC oncogene that is amplified in neuroblastoma cells and is common in small
 CC cell lung cancers. The present invention describes these cDNA molecules
 CC as useful for in hybridisation assays to detect expression of nucleic
 CC acids (or complementary nucleic acids) in a present in a given sample, as
 CC well as for screening assays by identifying molecules or compounds that
 CC specifically bind the cDNA as a ligand and modulate function or activity.
 CC Accordingly, these compositions exhibit cytostatic activity and can also
 CC be used for gene therapy purposes. This polynucleotide sequence is a cDNA
 CC that is differentially expressed in MYCN activated cells, given in an
 CC exemplification of the invention. NOTE: This sequence does not appear in
 CC the printed specification but has been obtained in electronic format from
 CC the US Patent Office at
 CC ftp.segdata.uspto.gov/sequence.html?DocID=20030119009.

XX Sequence 2216 BP; 593 A; 476 C; 524 G; 623 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.42e-167 Length: 2216
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 10 Gaps: 1

US-10-698-228-1 (1-351) x ADJ56536 (1-2216)

QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
 DB 189 CCTGAGCGGAGCCCGCTCCTGCGCAGACACCGGAGGAGATCTTCCAGGAGCCAC 248
 QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
 DB 249 GGAGCCGAAACTAAAGCAGCTGCCCGCGGTGGAGATGAGCCGCTGCTGAGAGAAA 308
 QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysG1 58
 DB 309 CCCCCCGCGCTTGTGCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 368
 QY 58 nAlaGlnAlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHistr 78
 DB 369 GGCAGAGGCTTCTTTTGGACCGCGGAGGAGGTGGACCTCTCCAAGGACATTCAGCACTG 428
 QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98
 DB 429 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCCATCTTCTGGCTTTCTTTCGAGC 488
 QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
 DB 489 AAGCATGGATAGTAATAATGAAAACCTTGGTGGACGATTTAGCCAGAGATTCAGATTAC 548
 QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
 DB 549 AGAAGCCCGCTGTTCTATGGCTTCCAAATTCGCAATGCGAAACATACATCTCTGAAATGTA 608
 QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
 DB 609 TAGTCTCTTTATTGACACTTACATAAAAGATCCCAAGAAAGAGGAATTTCTCTCAATGC 668
 QY 158 aIleGluThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrPheAlaAs 178
 DB 669 CATTGAAACCATGCTTGTGTCAAGAAAGAGGACAGACTGGCCCTTGCCTGGATTGGGA 728
 QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePh 198
 DB 729 CAAAGAGGCTACCTATGCTGAAACGCTGTTGTAGCCCTTTGCTCAGTGGAGGCAATTTCTT 788
 QY 198 eSerGlySerPheAlaAlaIlePheThrProLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 DB 789 TTCGGTCTCTTTTGGCTCGATATTTCTGGCTCAAGAAACGAGGACTGATGCTCGGCTCAC 848
 QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 DB 849 ATTTTCTAATGAACCTTATAGCAGATGAGGGTTTACACTGTGATTTGCTTGCCTGAT 908
 QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluLilLeValAspAl 258
 DB 909 GTTCAACACCTGGTACACAAACCATCGAGGAGAGAGATGAAGAGAAATAATTATCAATGC 968
 QY 258 aValLysIleGluGlnGluPheThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 DB 969 TGTTCGATAGAAACAGGAGTTCTCTCACTGAGGCCCTTGCCTGTAAGCTCATTTGGATGAA 1028
 QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluG1 298
 DB 1029 TTGCACCTTAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACTGGG 1088
 QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
 DB 1089 TTTTACAGGTTTTCAGAGTAGAGAACCCATTTGACTTATGGAGAAATATTTCACTGGA 1148

QY 318 uGlyLysThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
 DB 1149 AGGAAGACTAATCTTTTGAGAGAGAGTGGCGAGTATCAGAGGATGGGAGTGTATGTC 1208
 QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaPhe 351
 DB 1209 AAGTCAACAGAGAAATCTTTTACCTTGATGCTGACTTC 1248

RESULT 11

AA544917/c
 ID AAS44917 standard; DNA; 2482 BP.
 AC AAS44917;
 DT 18-DEC-2001 (first entry)
 DE Human contig polynucleotide sequence #170.
 KW Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
 KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
 KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
 KW nervous system disorder; inflammatory disorder; cell differentiation; de;
 KW angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
 KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;
 KW cytostatic; antirheumatic; antiarthritic; vulnery; antiinflammatory;
 KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
 KW neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;
 KW immunostimulant; analgesic; gene therapy.

OS Homo sapiens.
 OS Synthetic.

XX WO200164834-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004926.

XX 28-FEB-2000; 2000US-00515126.

XX 18-MAY-2000; 2000US-00577409.

XX 17-JUN-2000; 2000US-00597707.

XX 14-JUL-2000; 2000US-00616807.

XX 19-SEP-2000; 2000US-00664641.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QX, Ren F;

PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;

PI Drmanac R;

XX WPI; 2001-589862/66.

DR P-PSDB; AAU28017.

XX Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis, treatment of cancer,
 PT neurological, inflammatory disorders and for use in arrays for detection.
 PS Claim 1; SEQ ID NO 514; 153pp; English.

XX Sequences AAS44576-AAS44919 represent full-length polynucleotides and
 CC contig polynucleotides encoding polypeptides of the invention. The DNA
 CC and protein sequences are useful for the treatment, diagnosis and
 CC prevention of various types of disorder in a mammalian subject such as a
 CC human, dog, monkey, mouse, hamster or rat. The disorders include cancers
 CC such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such
 CC as multiple sclerosis, connective tissue diseases, rheumatoid arthritis,
 CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
 CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
 CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
 CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's
 CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
 CC bowel disease. The sequences exhibit activity relating to angiogenesis,

CC cell proliferation, cell differentiation, stem cell growth factor,
 CC activin or inhibin. Therefore, they can be used to manipulate stem cells
 CC in culture to give rise to neuroepithelial cells that can be used to
 CC augment or replace cells damaged by illness, accidental damage or genetic
 CC disorders. The sequences may also be used for regeneration of bone,
 CC cartilage, tendons and ligaments and in tissue repair and burn healing.
 CC Note: Some sequences for this patent did not form part of the printed
 CC specification, but were obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 2482 BP; 707 A; 598 C; 522 G; 655 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.68e-167 Length: 2482
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 4 Gaps: 1

US-10-698-228-1 (1-351) x AAS44917 (1-2482)

QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerAspThrAs 23
 DB 2181 CCTGAGCGGAGCCCGCGTCTCTGGCCAGCAAGACCGCGAGGAGGATCTTCCAGGAGCCAC 2122
 QY 23 nGluSerGluLeuLysSer-----AsnGluGluProLeuLeuArgLysSe 38
 DB 2121 GGAGCGCAAACTAAAGCAGCTGCCCGCGCGTGGAGGATGAGCGCTGTGAGAGAAA 2062
 QY rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysG1 58
 DB 2061 CCCCCCGCGCTTGTCTATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAGAA 2002
 QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAlaAspLeuSerLysAspLeuProHisTyr 78
 DB 2001 GGCAGAGGCTTCTCTTTTGGACCGCGCGAGGAGGTGACCTCTCCAAGGACATTCAGCAGCTG 1942
 QY 78 pAlenLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaA1 98
 DB 1941 GGAATCCCTGAAACCCGAGGAGAGATATTTTATATCCATGTTCTGGCTTTCTTTGCGAGC 1882
 QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
 DB 1881 AAGCGATGGCATAGTAAATGAAACCTTGGTGGAGCGGATTTAGCCCAAGATTCAGATTAC 1822
 QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
 DB 1821 AGBAGCCCGCTGTTCTATGGCTTCCAAATTCGCAATGCAAAACATACATCTCGAATGTA 1762
 QY 138 rSerLeuLeuLeuAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
 DB 1761 TAGTCTTCTTATTTGACACTTACATAAAAGATCCCAAAAGAAAGGAATTTCTTCTCAATGC 1702
 QY 158 alleGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
 DB 1701 CATTGAAACGATGCTTGTGTGTCGAGAGGAGGAGCTGGGCTTGGCTGATGTTGGGGA 1642
 QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
 DB 1641 CAAAGAGGTACTATGTTGTAAGCTGTTGTAGGCTTTGTGCTGAGTGGAGGCATTTTCTT 1582
 QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 DB 1581 TTCCGGTCTTTTTCGTCGATATTTCTGGCTCAAGAAACAGGAGGATGATGCTGCGCTCAC 1522
 QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 DB 1521 ATTTTCTAATGAACCTTATTAGCAGAGATGAGGGTTTACACTGTGATTTTGTCTCCCTGAT 1462
 QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
 DB 1461 GTTCAACACCTGGTATACAAACCAATCGGAGGAGAGATGAGAGAAATATTATCAATGC 1402


```

QY 258 avallysileGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
DB 1401 TGTTCGGATAGACAGGAGTCTCTACTGAGCGCTTGCTGTGAAGCTCATTTGGATGAA 1342
QY 278 nCysIleLeuMetLysGlnTyrlleGluPheValAlaAspArgLeuValGluLeuGl 298
DB 1341 TTGCACTCTAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACCTGGG 1282
QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
DB 1281 TTTTAGCAGAGGTTTTCAGAGTAGAGAACCATTTGACCTTTATGGAGAAATATTTCACTGGA 1222
QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrlGlnArgPheAlaValMetAl 338
DB 1221 AGGAAGACTAATCTCTTTGAGAAGAGAGTAGCGGAGTATCAGAGGATGGGAGTGTATGTC 1162
QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1161 AAGTCCAAACAGAGAATTTCTTTACCTTGGATGCTGACTTC 1122

RESULT 12
ABL65414
ID ABL65414 standard; DNA; 2500 BP.
XX
AC ABL65414;
XX
DT 15-MAY-2002 (first entry)
DE
DE Lung cancer related gene sequence SEQ ID NO:3751.
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; Gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 20-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.

```

```

PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX (AVAL-) AVALON PHARM.
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
DR WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
XX Claim 1; SEQ ID NO 3751; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 847 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms'
CC tumour
XX
XX Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 1,7e-167 Length: 2500
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservatives: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: Gaps: 1

US-10-698-228-1 (1-351) x ABL65414 (1-2500)
QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
DB 302 CCTGAGCGGAGCCCGGCTCTGCGCCAGACACCGCGAGGAGGATCTTCCAGAGGCCAC 361
QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
DB 362 GGAGCCGAAACTAAAGCAGCAGCTCCCGCGCGTGGAGGATGAGCGGCTGCTGAGAGAAA 421
QY 38 rSerArgArgPheValIlePheProIleGlnTyrlProAspIleTrpLysMetTrpLysGl 58
DB 422 CCCCCCGCGCTTGTCTATCTTCCCATCGAGTACCATGATATCTGGCAGATGATATAGAA 481
QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAlaAspLeuSerLysAspLeuProHisTr 78
DB 482 GGCAGAGGCTTCTCTTTTGGACCCCGCGAGGAGGTGACCTCTCCCAAGGACATTCAGCACTG 541

```


QY 78 pAenLysLeuLysAlaAspGluLysTyrPheLeSerHisIleLeuAlaPhePheAlaA 98
Db 542 CGAATCCCTGAAACCGAGAGAGATATTTATATCCATGCTTCTGGCTTTCTTCGAGC 601
QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 602 AAGCGATGCATAGTAAATGAAACCTTGTGGAGCGATTTAGCCAAAGAGTTCCAGATTAC 661
QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 662 AGAAGCCCGCTGTTCTATGGCTTCNAATTCATGGGAACATACATCTCGAANGTA 721
QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db 722 TAGTCTTCTATTGACACTTACATAAAGATCCCAAGAAAGGGAATTTCTCTTCAATGC 761
QY 158 alleGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
Db 782 CATTGAACGATGCTTGTGTCAAGAAAGAGGAGAGCTGGGCTTGGCTGGATTGGGA 841
QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
Db 842 CAAAGAGGTACTGTGGTGAACGCTTTGTAGCTTTGTGTCAGTGAAGGCAATTTCTT 901
QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db 902 TTCGGTCTTTTGGTGCATATTTCTGGCTCAAGAAACGAGGACTGATGCTGGCTCAC 961
QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db 962 ATTTCTTAATGAATATTATGACAGATGAGGGTTTACACTGTGATTTTGTCTGCCTGAT 1021
QY 238 tPheGlnTyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAl 258
Db 1022 GTTCAACACCTGGTACACAAACCATCGAGAGAGAGTAAAGAGAAATAATTATCAATGC 1081
QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db 1082 TGTTGGATAGAACAGAGTTCTACTAGGCTTGGCTGTGAAGCTCATTTGGAGTAA 1141
QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuG 298
Db 1142 TTGCACTTAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAATGGG 1201
QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG 318
Db 1202 TTTTAGCAAGGTTTTCAGAGTAGAAGAACCATTTGACTTTATGGAGAAATATTTCACTGGA 1261
QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
Db 1262 AGGAAGACTAATCTTTTGAGAGAGAGTAGGGAGTATCAGAGGATGGGAGTGATGTC 1321
QY 338 aGluThrTrpAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1322 AAGTCACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 13

ABL66517

ID ABL66517 standard; DNA; 2500 BP.

XX AC ABL66517;

XX AC ABL66517;

XX DT 15-MAY-2002 (first entry)

XX DE Lung cancer related gene sequence SEQ ID NO:4854.

XX DE Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;

XX DE stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;

XX DE cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;

XX DE gene; db.

XX OS Homo sapiens.

CC The data collected with respect to the anti-neoplastic agent as a result
CC of MI, and the data is sufficient to convey the chemical structure and/or
CC anti-neoplastic agent, and can be used for producing a product which is
CC activity and can be used in gene therapy. MI can be used for screening an
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC comprises a sequence (S) selected from 847 sequences (given in ABL61664
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC to be tested for anti-neoplastic activity, determining a change in
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC The present invention describes a method (M1) for screening for an anti-

PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.

PS Claim 1; SEQ ID NO 4854; 44pp; English.

XX The present invention describes a method (M1) for screening for an anti-
XX neoplastic agent. The method involves exposing cells to a chemical agent
XX to be tested for anti-neoplastic activity, determining a change in
XX expression of at least one gene (I) of a signature gene set, where (I)
XX comprises a sequence (S) selected from 847 sequences (given in ABL61664
XX to ABL70110), or is at least 95% identical to (S), where a change in
XX expression is indicative of anti-neoplastic activity. (I) has cytostatic
XX activity and can be used in gene therapy. MI can be used for screening an
XX anti-neoplastic agent, and can be used for producing a product which is
XX the data collected with respect to the anti-neoplastic agent as a result
XX of MI, and the data is sufficient to convey the chemical structure and/or

PA (AVAL-) AVALON PHARM.

PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;

PI Soppet DR, Weaver Z;

XX WPI; 2002-188264/24.

DR WPI; 2002-188264/24.

XX Screening for anti-neoplastic agent involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, and determining a change
XX in expression of a gene of a signature gene set.

CC properties of the agent. M1 can be used in the treatment of cancer such
 CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
 CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
 CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
 CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms
 CC tumour

SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,7e-167 Length: 2500
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 6 Gaps: 1

US-10-698-228-1 (1-351) x ABL6517 (1-2500)

QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
 DB 302 CTTGAGCGGACCCGGCTCTGCCAGCAGACCGGAGGAGGATCTTCAGGAGCCAC 361
 QY 23 nGluSerGluLeuLysSer-----AenGluGluProLeuLeuArgLysSe 38
 DB 362 GGAGCCGAAACCTAAAGCAGCTGCCCGCGCTGGAGGATGAGCCGCTCTCAGAGAAAA 421
 QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysG 58
 DB 422 CCCCCCGCGCTTGTTCATCTCCCATCGAGTACCATGATATCTGCAGATGTATAAGAA 481
 QY 58 nAlaGlnAlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
 DB 482 GGCAGAGGCTTCTTTGGACCGCCGAGGAGGTTGACCTCTCCAGACATTCAGGACG 541
 QY 78 pAsnLysLeuLysAlaAlaGlyLysTyrPheLysHisIleLeuAlaPhePheAlaAl 98
 DB 542 GGAATCCCTGAAACCGAGGAGAGATATTTATATCCCATGTTCTGGCTTCTTTCGAGC 601
 QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
 DB 602 AAGCGATGCGATAGTAAATGAAACCTTGGTGAGCGGATTTAGCCAAAGAGTTTCAGATTAC 661
 QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyr 138
 DB 662 AGAAGCCCGCTGTTCTATGGCTTCCAAATGCCATGGAAACATACATCTCGAATGTA 721
 QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPhePheAsnAl 158
 DB 722 TAGTCTCTTATGTACACTTACATAAAGATCCCAAGAAAGGGAATTTCTCTCAATGC 781
 QY 158 aIleGluThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAs 178
 DB 782 CATTGAAACGATGCTTGTGTCAAGAAGAGGAGGAGCTGGGCTTCGCTGGATGGGGA 841
 QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGluValPhePh 198
 DB 842 CAAAGAGGCTACTATGGTGAACGTTGTAGCTTGTGCTAGTGAAGGAGGCAATTTCTT 901
 QY 198 eSerGlySerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 DB 902 TTCGGTCTTTTGGTGCATATTTCTGGCTCAAGAAACGAGGACTGATGCTGGCTCAC 961
 QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 DB 962 ATTTTCTAATGAATATTATGACAGATGAGGGTTTACACTGTGATTTTGTCTGCTGAT 1021
 QY 238 tPheGlnTyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAl 258
 DB 1022 GTTCAACACCTGGTACACAAACCATTCGAGGAGGAGTAAAGAGAAATAATATCAATGC 1081
 QY 258 aValLysIleGluGlnGluPheLeuThrGlnAlaLeuProValGlyLeuIleGlyMetAs 278

DB 1082 TGTTCGATAGAACAGAGGTTCTCTCACTGAGGCTTGCCTGTGTGAAGCTATTGGGATGAA 1141
 QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG 298
 DB 1142 TTGCATCTTAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTTGGAACTGGG 1201
 QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG 318
 DB 1202 TTTTAGCAAGTTTTCAGAGTAGAGAACCCATTGACTTTATGAGAAATATTTCACCTGA 1261
 QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
 DB 1262 AGGAAAGACTTAATCTCTTTGAGAAGAGAGTAGGCGAGTATCAGAGGATGGGAGTGTATC 1321
 QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1322 AAGTCCACACAGAAATCTTTTACCTTGGATGCTGACTTC 1361
 RESULT 14
 ABL65859
 ID ABL65859 standard; DNA; 2500 BP.
 XX
 AC ABL65859;
 DX
 TT 15-MAY-2002 (first entry)
 XX
 DE Lung cancer related gene sequence SEQ ID NO:4196.
 XX
 KW Human; cancer; colon; breast; kidney; oesophagus; thyroid; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilms's tumour; adenocarcinoma;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200194629-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US010838.
 XX
 PR 05-JUN-2000; 2000US-0209473P.
 PR 05-JUN-2000; 2000US-0209531P.
 PR 18-SEP-2000; 2000US-0233133P.
 PR 18-SEP-2000; 2000US-0233617P.
 PR 20-SEP-2000; 2000US-0234009P.
 PR 20-SEP-2000; 2000US-0234034P.
 PR 20-SEP-2000; 2000US-0234052P.
 PR 22-SEP-2000; 2000US-0234509P.
 PR 22-SEP-2000; 2000US-0234567P.
 PR 25-SEP-2000; 2000US-0234923P.
 PR 25-SEP-2000; 2000US-0234924P.
 PR 25-SEP-2000; 2000US-0235077P.
 PR 25-SEP-2000; 2000US-0235082P.
 PR 25-SEP-2000; 2000US-0235134P.
 PR 25-SEP-2000; 2000US-0235280P.
 PR 26-SEP-2000; 2000US-0235637P.
 PR 26-SEP-2000; 2000US-0235638P.
 PR 27-SEP-2000; 2000US-0235711P.
 PR 27-SEP-2000; 2000US-0235720P.
 PR 27-SEP-2000; 2000US-0235840P.
 PR 27-SEP-2000; 2000US-0235863P.
 PR 28-SEP-2000; 2000US-0236028P.
 PR 28-SEP-2000; 2000US-0236032P.
 PR 28-SEP-2000; 2000US-0236033P.
 PR 28-SEP-2000; 2000US-0236034P.
 PR 28-SEP-2000; 2000US-0236109P.
 PR 28-SEP-2000; 2000US-0236111P.
 PR 29-SEP-2000; 2000US-0236842P.
 PR 29-SEP-2000; 2000US-0236891P.
 PR 02-OCT-2000; 2000US-0237172P.
 PR 02-OCT-2000; 2000US-0237173P.
 PR 02-OCT-2000; 2000US-0237278P.

PR 02-OCT-2000; 2000US-0237294P.
 PR 02-OCT-2000; 2000US-0237295P.
 PR 02-OCT-2000; 2000US-0237316P.
 PR 03-OCT-2000; 2000US-0237425P.
 PR 03-OCT-2000; 2000US-0237598P.
 PR 03-OCT-2000; 2000US-0237604P.
 PR 03-OCT-2000; 2000US-0237606P.
 PR 03-OCT-2000; 2000US-0237608P.
 PR 01-NOV-2000; 2000US-0244867P.
 PR 01-NOV-2000; 2000US-0245084P.
 XX (AVAL-) AVALON PHARM.
 XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX WPI; 2002-188264/24.
 XX
 PT Screening for anti-neoplastic agent involves exposing cells to a chemical
 PT agent to be tested for anti-neoplastic activity, and determining a change
 PT in expression of a gene of a signature gene set.
 XX
 PS Claim 1; SEQ ID NO 4196; 44pp; English.
 XX
 CC The present invention describes a method (M1) for screening for an anti-
 CC neoplastic agent. The method involves exposing cells to a chemical agent
 CC to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening an
 CC anti-neoplastic agent, and can be used for producing a product which is
 CC the data collected with respect to the anti-neoplastic agent as a result
 CC of M1, and the data is sufficient to convey the chemical structure and/or
 CC properties of the agent. M1 can be used in the treatment of cancer: such
 CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
 CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
 CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
 CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
 CC tumour
 XX
 SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 1.7e-167 Length: 2500
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.23% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 6 Gaps: 1
 US-10-698-228-1 (1-351) x ABL65859 (1-2500)
 QY 4 ProGluArgProGluAlaAlaGlyLeuAspGluArgSerSerSerAspThrAs 23
 DB 302 CCGAGCGGAGCCCGCGTCTGGCCAGCAAGACCGCGAGGAGGATCTTCAGAGGCCAC 361
 QY 23 nGluSerGluLeuLysSer-----AanGluGluProLeuLeuArgLysSe 38
 DB 362 GGAGCGGAAACTAAAGCAGCTGCCCGCGGTGGAGGATGAGCCGCTCTGAGAGAAA 421
 QY 38 rSerArgArgPheValilePheProileGlnTyrProAspileTrpLysMetTyrLysG1 58
 DB 422 CCCC CGCGCTTTGTCATCTCCCATCGAGTACCATGATATCTGCAGATGTATAAGAA 481
 QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
 DB 482 GCGAGAGGCTTCCTTTTGACCGCGCGAGGAGTTGACCTCTCCAGAGCATTCAGCACTG 541
 QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheTleSerHisleLeuAlaPhePheAla 98
 DB 542 GGAATCCCTGAAACCGGAGGAGAGATATTTATATCCCAATGTCGTCTTCTTTCAGC 601

QY 98 aserAspGlylleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
 DB 602 AAGCGATGGCATAGTAATAAGAAACTTGTGGAGCGATTAGCCAAAGATTCAGATTAC 661
 QY 118 oGluAlaArgCysPheTyrGlyPheGlnileLeuLeileGluAsnValHisSerGluMetTy 138
 DB 662 AGAAGCCCGCTGTTTCTATGGCTTCCAAATTTGCCATTGCCAAGAAACATACATTCTGAAATGTA 721
 QY 138 rSerLeuLeuileAspThrTyrileArgAspProLysLysArgGluPheLeuPheAsnAl 158
 DB 722 TAGTCTCTTATTGACACTTACATAAAGATCCCAAGAAAGGGAATTTCTTCAATGC 781
 QY 158 alileGluThrMetProTyrVallysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
 DB 782 CATTAACACGATGCTTGTCTCAAGAAAGAGGAGACTGGGCTTGGCCTGATGGGGA 841
 QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
 DB 842 CAAAGAGGCTACCTATGTTGTAACGTTGTAGCCTTTGCTGAGTGGAGGCAATTTCTT 901
 QY 198 eSerGlySerPheAlaAlailePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 DB 902 TTCGGGTCTTTTGGCTGCGATATTCTGGCTCAAGAAACGAGGACTGATGCTGGCTCAC 961
 QY 218 rPheSerAsnGluLeuileSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 DB 962 ATTTTCTTAATGAATTTATTAGCAGAGATGAGGGTTTACACTGTGATTTTGTGCTGAT 1021
 QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluLeuileValAspAl 258
 DB 1022 GTTCAACACCTGGTACACAAACCATCGGAGGAGAGTAGAGAAATAATATCAATGC 1081
 QY 258 aValLysilleGluGlnGlnPheLeuThrGluAlaLeuProValGlyLeuileGlyMetAs 278
 DB 1082 TGTTCGATAGAACAGAGTTCTCTCACTGAGGCTTCTGCTGTAAGCTCATTTGGATGAA 1141
 QY 278 nCysilleLeuMetLysGlnTyrileGluPheValAlaAspArgLeuLeuValGluG1 298
 DB 1142 TTGCACCTTAATGAAGCAATACATTTGATTTGTGGCAGACAGACTTATGCTGGAACTGG 1201
 QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnileSerLeuG1 318
 DB 1202 TTTTAGCAAGGTTTTCAGATAGAACCCCATTTGACTTTATGGAGAAATATTTCACTGGA 1261
 QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
 DB 1262 AGCAAGACTAACTTCTTTGAGAAGACAGATAGGCGAGTATCAGAGGATGGGAGTGATCTC 1321
 QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1322 AAGTCCAACAGAGAAATTTCTTTTACCTTGGATGCTGACTTC 1361
 RESULT 15
 ABX10335
 ID ABX10335 standard; DNA; 2500 BP.
 XX
 AC ABX10335;
 XX
 DT 28-JAN-2003 (first entry)
 XX
 DE DNA encoding protein differentially regulated in prostate cancer #4.
 XX
 KW Prostate cancer; gene expression; differential regulation;
 KW molecular marker; drug target; cancer detection; cancer diagnosis;
 KW cancer staging; cancer grading; cancer assessing; cancer monitoring;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200281638-A2.
 XX
 PD 17-OCT-2002.

XX 08-APR-2002; 2002WO-US010824.
 XX 06-APR-2001; 2001US-0281731P.
 PR 06-APR-2001; 2001US-0281732P.
 XX (ORIG-) ORIGENE TECHNOLOGIES INC.
 XX PA
 XX PI Sun Z, Jay G;
 XX WPI; 2003-058520/05.
 DR P-PSDB; ABU07433.
 XX
 PT Novel genes which are differentially regulated in prostate cancer, useful
 PT for diagnosing prostate cancer in prostate tissue sample and assessing
 PT therapeutic or preventive intervention in prostate cancer patients.
 XX
 PS Claim 1; Page 71-72; 416pp; English.
 CC The invention describes genes (I) which are differentially regulated in
 CC prostate cancer. (I) Is useful for diagnosing a prostate cancer in a
 CC sample comprising prostate tissue, which involves determining the number
 CC of target genes which are differentially-regulated in the sample, where
 CC the number is indicative of the probability that the sample comprises
 CC prostate cancer. (I) Is useful for assessing a therapeutic or preventive
 CC intervention in a subject having a prostate cancer, which involves
 CC determining the expression levels in a sample comprising prostate tissue
 CC of target genes which are differentially-regulated in prostate cancer.
 CC Preferably, the expression levels of at least 10 genes are determined.
 CC (I) is also useful for identifying agents that modulate a biological
 CC activity of a polypeptide differentially-regulated in prostate cancer
 CC cells, which involves contacting a polypeptide differentially-regulated
 CC in prostate cancer cells with a test agent under conditions effective for
 CC the test agent to modulate a biological activity of the polypeptide, and
 CC determining whether the test agent modulates the biological activity. (I)
 CC is useful as molecular markers, as drug targets, and for detecting,
 CC diagnosing, staging, grading, assessing, monitoring, prognosticating,
 CC preventing or treating, determining predisposition to diseases and
 CC conditions especially relating to prostate cancer. (I) and its expression
 CC products are used in the diagnostic test to assay for presence of cancer
 CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
 CC blood etc. (I) is useful for assessing cancer e.g., to determine the type
 CC of cancer, its stage of development, the nature of genetic defect, etc.
 CC The polypeptide encoded by (I) can be used as target for therapy or drug
 CC discovery. (I) can also be used for expressing the polypeptide and thus
 CC for searching specific binding partners of the polypeptide. (I) is useful
 CC in therapeutic applications to treat prostate cancer. The identification
 CC of specific genes, and groups of genes, expressed in pathways
 CC physiologically relevant to prostate cancer permits the definition of
 CC functional and disease pathways and the delineation of targets in these
 CC pathways which are useful in diagnostic, therapeutic, and clinical
 CC applications. This sequence encodes a protein differentially regulated in
 CC prostate cancer
 XX
 SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.7e-167 Length: 2500
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 8 Gaps: 1

US-10-698-228-1 (1-351) x ABX10335 (1-2500)

QY 4 ProGluArgProGluAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
 DB 302 CCGAGCGGAGCCCGGCTCTCGCCAGCAGACCGGAGGAGATCTTCAGGAGCCAC 361
 QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
 DB 362 GGAGCGGAAACTAAAGCAGCTGCCCGCGCTGGAGGATGAGCCGCTCTGAGAGAAA 421

QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysG1 58
 DB 422 CCCCCCGCGCTTTGTCTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGATATAGAA 481
 QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLysSerLysLeuProHisTr 78
 DB 482 GGCAGAGGCTTCTCTTTGGACCGCGAGAGGTTGACCTCTCCAAGGACATTCAGCATG 541
 QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98
 DB 542 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCATGTTCTTGGCTTCTTTGACG 601
 QY 98 aSerAspGlyIleValAsnGlnLeuValGluArgPheSerGlnGluValGlnValPr 118
 DB 602 AAGCATGGATAGTAATAAGAACTTGGTGGAGGATTTAGCCAGAGTTCAGATTAC 661
 QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
 DB 662 AGAAGCCCGCTGTTCTATGGCTTCCAAATTCGCATGGAACATACATTCGAAATGTA 721
 QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
 DB 722 TAGTCTCTTATTGACACTTACATAAAGATCCCAAGAAAGGGAATTTCTTCAATGC 781
 QY 158 aIleGluThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrIleAlaAs 178
 DB 782 CATTGAACCATGCTGCTCAAGAAAGAGGAGACCTGGCCCTTGGCTGGATGGGA 841
 QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
 DB 842 CAAAGAGGCTACCTGTTGCTGCAACGTTGTAGCTTTGCTGCAGTGAAGCAATTTCTT 901
 QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 DB 902 TTCGGTCTCTTTTCGCTCGATATTTCTGGCTCAAGAAACGAGGACTGATGCTGGCCTCAC 961
 QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 DB 962 ATTTCTAATGAACCTATTAGCAGAGATGAGGTTTACCTGTGATTTTGTCTGCTGAT 1021
 QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
 DB 1022 GTTCAAAACACCTGTCACAAACCATCGAGGAGAGAGTAAGAGAAATAATTATCAATGC 1081
 QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 DB 1082 TGTTCGATAGAACAGGAGTTCTCTACTGAGGCTTGCCTGTGAGAGCTCATTTGGATGAA 1141
 QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluG1 298
 DB 1142 TTCACCTCTAATGAACCAATACATTTGTTGGCAGACAGACTTATGCTGGAACTGGG 1201
 QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
 DB 1202 TTTTAGCAGGTTTTCAGATAGAGAACCATTTGACTTTATGGAGATAATTTTCACTGA 1261
 QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
 DB 1262 AGGAAAGACTTAATCTTTTGAAGAAGAGTAGGCGAGTATCAGAGGATGGGAGTGATGTC 1321
 QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1322 AGTCCAACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361

Search completed: October 30, 2005, 07:11:59
 Job time : 466 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 30, 2005, 07:01:57 ; Search time 2584 Seconds

(without alignments)
5170.495 Million cell updates/sec

Title: US-10-698-228-1

Perfect score: 1821

Sequence: 1 MGDPERPEAAGLDQDERSSS.....QRFVAVMAETTDNVFLDADF 351

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US10698228/runat.26102005.100715.5117/app_query.fasta.1.519
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228@cgn.1.1.2607@runat.26102005.100715.5117 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	100.0	3292	3 CR617553	CR617553 full-leng
2	1821	100.0	4748	3 HSM802033	AL137348 Homo sapi
3	1791	98.4	4650	3 BC042948	BC042948 Homo sapi
4	1737.5	95.4	3501	3 CR627376	CR627376 Homo sapi
5	1591	87.4	1083	1 AL547501	AL547501 AL547501
6	1585	87.0	1056	9 AY398973	AY398973 Homo sapi
7	1468	80.6	1043	9 AY398975	AY398975 Mus muscu
8	1460.5	80.2	1571	3 CR603461	CR603461 full-leng
9	1460.5	80.2	1573	3 CR625489	CR625489 full-leng

10	1460.5	80.2	1582	3	CR608076	full-leng
11	1460.5	80.2	1588	3	CR602054	full-leng
12	1460.5	80.2	1592	3	CR621427	full-leng
13	1460.5	80.2	1600	3	CR604378	full-leng
14	1460.5	80.2	1605	3	CR590959	full-leng
15	1460.5	80.2	1612	3	CR614990	full-leng
16	1460.5	80.2	1613	3	CR609838	full-leng
17	1460.5	80.2	1623	3	CR618451	full-leng
18	1460.5	80.2	1630	3	CR602150	full-leng
19	1460.5	80.2	1796	3	CR596700	full-leng
20	1443	79.2	2088	3	AK088907	Mus muscu
21	1384	76.0	879	5	BQ441857	AGENCOURT
22	1354	74.4	892	7	CN163214	952695 MA
23	1328.5	73.0	997	4	BM468712	AGENCOURT
24	1316	72.3	1008	9	AY398974	Pan trogl
25	1277	70.1	932	5	BUI96941	AGENCOURT
26	1275	70.0	914	7	CF995079	AGENCOURT
27	1257.5	69.1	946	7	CN024619	AGENCOURT
28	1235	67.8	870	7	CO648380	ILLUMIGEN
29	1222	67.1	836	7	CR416843	CR416843
30	1221	67.1	893	5	BUI90680	AGENCOURT
31	1209.5	66.4	923	7	CR580780	CR580780
32	1209	66.4	1076	4	BM460735	AGENCOURT
33	1203	66.1	801	5	BP680185	BP680185
34	1197.5	65.8	933	6	CA983633	AGENCOURT
35	1195	65.6	1111	5	BM914217	AGENCOURT
36	1189	65.3	909	6	CA981614	AGENCOURT
37	1188	65.2	804	6	CD656661	AGENCOURT
38	1186.5	65.2	1028	5	BQ050629	AGENCOURT
39	1186	65.1	879	5	BUI89053	AGENCOURT
40	1186	65.1	910	6	CD794331	EST665692
41	1177	64.6	909	7	CK179258	EST768578
42	1170	64.3	928	4	BM018808	603646723
43	1168.5	64.2	1080	6	CD461172	Fg09 04p0
44	1167	64.1	896	5	BUI70979	AGENCOURT
45	1164	63.9	838	5	BUI910710	AGENCOURT

ALIGNMENTS

RESULT 1

LOCUS CR617553 3292 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DI011F14 of Placentia Cot 25-normalized of Homo sapiens (human).
ACCESSION CR617553
VERSION CR617553.1 GI:50498360
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3292)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue
REFERENCE 2 (bases 1 to 3292)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with NotI and cloned into the NotI and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life technologies, a division of Invitrogen.
FEATURES
source Location/Qualifiers
1..3292 /organism="Homo sapiens"

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1011YF14"
/tissue_type="placenta Cot 25-normalized"
/plasmid="pCWSport_6"

ORIGIN

Alignment Scores:

Pred. No.: 2,61e-210 Length: 3292
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-698-228-1 (1-351) x CR617553 (1-3292)

```

QY 1 MetGlyAspProGluArgProGluAlaLaglyLeuAspGlnAspGluArgSerSer 20
Db 27 ATGGGCGACCCGGAAGCGCGAAGCGCGCGCTGGATCAGATCAGAGATCATCTTCA 86
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
Db 87 GACACCAACGAAGTGAATAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 146
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
Db 147 CGTTTGTTCATCTTCCATCCAGTACCTCTCATCTCTCATCTCTTTCGACGCCAGTGCAT 206
QY 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
Db 207 GCTTCTCTTCGACAGCAGAGAGGTGCGACTTATCAAGAGATCTCCTCACTGGACACAG 266
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
Db 267 CTTAAAGCAGATGAGAAGTACTCTCTCTCATCTCTAGCCCTTTTTCGACGCCAGTGCAT 326
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 327 GGAATGTAAATGAATAATTTGTGTGAGCGCTTTAGTCAGAGGTGCGAGGTTCCAGAGGCT 386
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 387 CGCTCTTTCTATGGCTTTCAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 446
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 447 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 506
QY 161 ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
Db 507 ACCATGCCCTATGTTAAGAAAGAGAGATGGGCTTTCGAGGATACGATAGAGAA 566
QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGluValPhePheSerGly 200
Db 567 TCTACTTTTGGGAAGAGTGTGTGGCTTTGTCTGTAGAGAGGATTTCTTCTCAGGA 626
QY 201 SerPheAlaAlaIlePheTyrLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 627 TCTTTTGTCTGTATATCTGGCTTAAAGAGAGAGGCTTTATGCCAGGACTCTCTTTTCC 686
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 687 AATGAACCTCATCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGTGCTGTATGTTCCAA 746
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 747 TACTTAGTAAATTAAGCTTCAGAGAAAGGGTCAGGAGATCATTTGTTGATGCTGTCAAA 806
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 807 ATTGACGAGGAGTTTAAACAGAGCCTTGCCAGTGTGGCTCATTTGGAATGAATGCATT 866

```

```

QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 867 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTGGATTCTCA 926
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 927 AAGGTTTTTTCAGGCAGAAAAATCCTTTTGTATTTTATGGAACAATTTCTTTAGAAAGGAAA 986
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 987 ACAAATTTTCTTTGAGAAACAGATTTTTCAGAGTATCAGCGTTTTCAGGTATGCGAGAAACC 1046
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1047 ACAGATAACGTTCTTTCACCTTGGATGCGAGATTTT 1079

RESULT 2
HSM802033      4748 bp      mRNA      linear      HTC 22-SEP-2004
LOCUS          Homo sapiens mRNA; cDNA DKFZp761E1312 (from clone DKFZp761E1312).
DEFINITION     AL137348
ACCESSION      AL137348.1 GI:6807859
VERSION        HTC.
KEYWORDS       Homo sapiens (human)
SOURCE          Homo sapiens
ORGANISM       Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                1 (bases 1 to 4748)
                Ansorge,W., Krieger,S., Regiert,T., Rittmuller,C., Schwager,B.,
                Wieses,H.W., Weill,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
                Wiemann,S.
                The German cDNA Consortium
CONSTRM        Direct Submission
TITLE          Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
JOURNAL        Neuherberg, GERMANY
COMMENT        Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
                Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
                sequenced by EMBL (European Molecular Biology Laboratories,
                Heidelberg/Germany) within the cDNA sequencing consortium of the
                German Genome Project.
                This clone (DKFZp761E1312) is available at the RZPD Deutsches
                Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
                Please contact RZPD for ordering:
                http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp761E1312
                Further information about the clone and the sequencing project is
                available at http://mips.gsf.de/projects/cdna/.
FEATURES        Location/Qualifiers
                1..4748
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="RZPD:DKFZp761E1312"
                /db_xref="taxon:9606"
                /clone="DKFZp761E1312"
                /tissue_type="amygdala"
                /clone_lib="761 (synonym: hamy2). Vector pSport1; host
                DH10B; sites NotI + SalI"
                /dev_stage="adult"
                /note="ribonucleotide reductase"
                1..4748
                /gene="DKFZp761E1312"
                47..1102
                /gene="DKFZp761E1312"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="CAB70703.2"
                /db_xref="GI:52545612"
                /db_xref="GOA:Q9NTD8"
                /db_xref="UniProt/TREMBL:Q9NTD8"
                /translation="MGDPPEAAGLDQDERSSDTNEISIKSNEEPLLRSSRRFVI
                FPIQPDIDWKYKQAQSFATFEVDLSKDLPHNKLKADEKFI SHILAFPAASDGI
                VNIENLVERFSQSVQVPEARCFYGFQILINHVHEMYSLIDTVIRDPKKEFLPNATE
                TWPYVKKADWALRWIADRKSTFGERVAPAAVEGVFSGSFAAIFWLKRGGLMPGLT
                FSNELISRDEGLHCDFAFLMFQYLVNKPSEVRREIIVDAVKIEQEFLEALPVGLIG

```


MNCILMKQVIEFVADRLLVELGFSKVFOAENPDPFMENISLBEGKTNFFPKRVSEYQRF
AVMAETTDNVFTLDAF"

ORIGIN

Alignment Scores:

Pred. No.: 4,56e-210 Length: 4748
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-698-228-1 (1-351) x HSM802033 (1-4748)

QY 1 MetGluValProGluArgProGluAlaAlaGlyLeuAspGluAspGluArgSerSer 20
DB 47 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGATCATCTTCA 106
QY 21 AspThrAsnGluSerGluLeuLeuYssSerAsnGluGluProLeuLeuArgLyssSerSerArg 40
DB 107 GACACCAACGAAGTGAATAAAGTCAATGAAGACCACTCTTAAGAAGAGTTCTCGC 166
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPlyMetTyrLyssGlnAlaGln 60
DB 167 CGGTTTGTCTATTTCCATCCAGTACCTCGATATTTGGAATGTATAAACAGGCACAG 226
QY 61 AlaSerPheThrAlaGluGluValAspLeuSerLyssAspLeuProHisTyrAsnLyss 80
DB 227 GCTTCTCTCTGACAGCAGAGAGGTGCGACTTATCAAGAGATCTCCCTCACTGGAACAAG 286
QY 81 LeuLyssAlaAspGluLyssTyrPheIleSerHisIleLeuAlaPhePheAlaLaserAsp 100
DB 287 CTTAAGCAGATGAGAGTACTTCTCTCACTCTGAGCTTTTTCGAGCAGCATGAT 346
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 347 CGAATTTGAATGAAATTTGGTGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 406
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 407 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 466
QY 141 LeuIleAspThrTyrIleArgAspProLyssLyssArgGluPheLeuPheAsnAlaIleGlu 160
DB 467 CTGATAGACACTTATCATAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 526
QY 161 ThrMetProTyrValLyssLyssAlaAspTyrAlaLeuArgTyrIleAlaAspArgLyss 180
DB 527 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTCGATGGATAGCAGATAGAAA 586
QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
DB 587 TCTACTTTTGGGGAAGAGTGTGGCTTTGCTGTGTAGAAGAGTCTTCTCTCAGCA 646
QY 201 SerPheAlaAlaIlePheTyrPleuLyssLyssArgGlyLeuMetProGlyLeuThrPheSer 220
DB 647 TCTTTTGTCTATATTTCTGGCTAAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC 706
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 707 ATGAACTCATCAGCAGATGAGAGCTTCACTGTGACTTTGCTGCTGCTGATTTCCAA 766
QY 241 TyrLeuValAsnLyssProSerGluGluArgValArgGluIleValAspAlaValLyss 260
DB 767 TACTTTAGTAAATAAGCCTTCAGAGAAGGGTCAAGGAGATCATTTGTCATGTCGCAAA 826
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 827 ATTGAGCAGAGTATTTTAAACAGAGCCTTGCCAGTTGGCTCATTTGGAATGAATTCATT 886
QY 281 LeuMetLyssGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 887 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTTACTTTGTGAATTTGATTTCTCA 946

QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 947 AAGGTTTTTCAGGCAGAAATCCTTTTGATTTTATGGAAACATTTCTTTAGAGGAAA 1006
QY 321 ThrAsnPhePheGluLyssArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 1007 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAGAAC 1066
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1067 ACAGATAACGCTCTTCCCTTGGATGCGAGATTTT 1099

RESULT 3
BC042948
LOCUS
DEFINITION
Homo sapiens cDNA clone IMAGE:4798175, containing frame-shift errors.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT

BC042948 4650 bp mRNA linear HTC 12-OCT-2004
Homo sapiens cDNA clone IMAGE:4798175, containing frame-shift errors.
BC042948 GI:34194000
HTC.
Homo sapiens (human)
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 4650)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,P., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schectz,T.E., Brownstein,M.J., Usdin,T.B., Toshnyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalek,U., Smallos,D.E., Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 4650)
Director MGC Project.
Direct Submission
Submitted (09-JAN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 25, 2003 this sequence version replaced gi:27695575.
Contact: MGC help desk
Email: cgapsb@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshioyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabhu, Parvaneh Saeedi, JR Santos, Angeliue Scherch, Ursula Skalska, Duane Smallos, Jeff Stott, Miranda Tsai, George Yang, Jacquie

Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E.B. Consortium/BLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 70 Row: h Column: 19
 This clone has the following problem: frame shifted.

FEATURES

source

```
1. .4650
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4798175"
/tissue_type="Brain, hypothalamus"
/clone_lib="NIH MGC_96"
/lab_host="DH10B"
/notes="Vector: pBluescript"
```

ORIGIN

Alignment Scores:

```
Pred. No.: 2,04e-206 Length: 4650
Score: 1791.00 Matches: 350
Percent Similarity: 99.72% Conservative: 0
Best Local Similarity: 99.72% Mismatches: 1
Query Match: 98.35% Indels: 1
DB: 3 Gaps: 0
```

US-10-698-228-1 (1-351) x BC042948 (1-4650)

```
QY 1 MetGlyAspProGluArgProGluAlaLaglyLeuAspGlnAspGluArgSerSer 20
DB 72 ATGGGCGACCCGGAAGCGCGAAGCGCGCGGCTGGATCAGATGAGATCATCTTCA 131
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 132 GACACCAACGAAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 190
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
DB 191 CGGTTTGTCATCTTCCATCCAGTACCTCTGATATTTGGAAATGATATAACAGGCACAG 250
QY 61 AlaSerPheTyrTrpAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
DB 251 GCTTCTCTCTGCACAGCAGAGAGTGCAGTCTATCAAGGATCTCTCTCACTGGAACAAG 310
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaLalaSerAsp 100
DB 311 CTTAAAGCAGATGAGAGTACTTCTCTCTCACTCTTACCTTTTTCAGCAGTGTAT 370
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 371 GGAATGTAAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGAGTGCAGGTTCCAGAGGCT 430
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 431 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 490
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 491 CTGATAGACACTTACATCAGATATCCAAAGAAAGGGAATTTTATTAATGAATGA 550
QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
DB 551 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTTCGATGGATAGCAGATAGAAA 610
QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
DB 611 TCTACTTTTGGGGAAGAGTGGTGGCTTCTGCTGTAGAGAGAGTCTTCTCTCAGGA 670
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 671 TCTTTTGCCTGCTATATCTGGCTAAAGAAAGAGAGGCTTATGCGCAGAGCTCACTTTTCC 730
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
```

```
DB 731 AATGAACATCATCAGCAGATGAAGACTTCACTGTGACTTTGTGCTGTGCTGTGCTCA 790
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleLeuValAspAlaValLys 260
DB 791 TACTTGTAAATAAGCTTTCAGAAAGAAAGGTCAGGAGATCATTTGTGTGCTGTGCTCA 850
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 851 ATTGAGCAGGAGTTTAAACAGAGCTTCCAGTTGGCTCATTTGGATGAATGATTCATT 910
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 911 TTGATGAAACAGTACATTTAGTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTC 970
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 971 AAGGTTTTTCAGGCAGAAAAATCTTTTGAATTTTATGAAAAACATTTCTTTTGAAGGAAA 1030
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 1031 ACAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGGTTTTCAGTTATGCGAGAAACC 1090
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1091 ACAGATAACGCTCTTCACTTTGGATGCGAGATTTT 1123

RESULT 4
CR627376 3501 bp mRNA linear HTC 22-SEP-2004
LOCUS Homo sapiens mRNA; cDNA DKFZp686M05248 (from clone DKFZp686M05248) .
DEFINITION CR627376
ACCESSION CR627376.1 GI:50949847
VERSION HTC.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3501)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Ostenwaeider,B., Obermaier,B., Deutschenbaur,S., Schaiipp,A.,
Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
The German cDNA Consortium
Direct Submission
Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Medigenomix (Martinsried/Germany) within the cDNA
sequencing consortium of the German Genome Project.
This clone (DKFZp686M05248) is available at the RZPD Deutsches
Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp686M05248
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES
source
1. .3501
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686M05248"
/tissue_type="salivary gland"
/lab_host="DH10B; sites Sf1A + Sf1B"
/dev stage="adult"
/notes="ribonucleotide reductase M2 B (TP53 inducible)"
1. .3501
/gene="DKFZp686M05248"
332. .1225
/gene="DKFZp686M05248"
/codon_start=1
/product="hypothetical protein"

gene
CDS
```



```
Db 27 ATGGCGCGCCGGAAGGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTC 86
Qy 20 rAspThrAnGluSerGluIleLysSerAnGluProLeuLeuAArgLysSerSerAr 40
Db 87 AGACACCAACGAAAGTGAATAAGTCAATGAAGAGCCACCTCTAAGAAAGAGTTCTCG 146
Qy 40 gArgPheValIlePheProIleGlnTrpProAspIleTrpLysMetTrpLysGlnAlaGl 60
Db 147 CCGGTTTGTCATCTTCCANTCAGTACCCTGATATTTGGAATGTAATAACAGGCACA 206
Qy 60 nAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAnLys 80
Db 207 GCCTTCCTTCGCAGCAGAGAGCTCGACTATCAAGAGTCTCCCTCAGTGGAAACA 266
Qy 80 sLeuLysAlaAspGluLysTrpPheIleSerHisIleLeuAlaPhePheAlaAlaSerAs 100
Db 267 GCTTAAAGCAGATGAGAAGTACTTCTCTCACATCTTGAAGCTTTTTCAGCAGCAGTGA 326
Qy 100 pGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAl 120
Db 327 TGGAAATGTGAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTCCAGGTCCAGAGGC 386
Qy 120 aArgCysPheTrpGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTrpSerLe 140
Db 387 TCGCTGTTCTATGGCTTCAATCTCATCGAATGTTCACTCAGAGATGTACAGTTT 446
Qy 140 uLeuIleAspThrTrpIleAArgAspProLysLysArgGluPheLeuPheAsnAlaIleGl 160
Db 447 GCTGATAGACATTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGA 506
Qy 160 uThrMetProTrpValLysLysIleAlaAspTrpAlaLeuArgTrpIleAlaAspArgLy 180
Db 507 AACCATGCTCTATGTTAAGAAAGAGAGATGGGCTTGGCATGATGATGAGATGAGAA 566
Qy 180 sSerThrPheGlyGluArgValAlaPheAlaAlaValGluGluValPhePheSerGl 200
Db 567 ATCTACTTTTGGGAAAGAGTGGTGGCTTGTCTGTAGAAGGAGTTTCTTCTCAGG 626
Qy 200 ySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSe 220
Db 627 ATCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTC 686
Qy 220 rAnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGl 240
Db 687 CAATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTGTGCTGATGTCCA 746
Qy 240 nTrpLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLy 260
Db 747 ATACTTAGTAATTAAGCTTTCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAA 806
Qy 260 sIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIl 280
Db 807 AATTGAGCAGGAGTTTAAACAGAGCCTTGCCAGTTGCCCTCATTTGGAATGAATTCAT 866
Qy 280 eLeuMetLysGlnTrpIleGluPheValAlaAspArgLeu-LeuValGluLeuGlyPheS 300
Db 867 TTTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAMCTTGGATTCT 926
Qy 300 eLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyL 320
Db 927 CAAGAGKTTTTCAGGCAGAAATTCCTTTGATTTTATGGAAM-AATTCTTTAAGAGGAA 985
Qy 320 yEThrAsnPhePhe 324
Db 986 AAMA-AATTTCTTT 998
```

RESULT 6

AY398973

LOCUS

DEFINITION

Homo sapiens HCM0069 gene, VIRTUAL TRANSCRIPT, partial sequence,

genomic survey sequence.

ACCESSION

AY398973

```
VERSION AY398973.1 GI:39754962
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1056)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civeello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1056)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civeello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
    source 1..1056
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        <1..>1056
        /locus_tag="HCM0069"
gene
    gene 1..1056
        /locus_tag="HCM0069"
ORIGIN
Alignment Scores:
Pred. No.: 3,02e-182 Length: 1056
Score: 1595.00 Matches: 313
Percent Similarity: 89.17% Conservative: 0
Best Local Similarity: 89.17% Mismatches: 38
Query Match: 87.04% Indels: 0
DB: 9 Gaps: 0
US-10-698-228-1 (1-351) x AY398973 (1-1056)
```

```
Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
Db 1 ATGGCGCGCCGGAAGGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTC 60
Qy 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuAArgLysSerSerArg 40
Db 61 GACACCAACCAAGTGAATGAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Qy 41 ArgPheValIlePheProIleGlnTrpProAspIleTrpLysMetTrpLysGlnAlaGln 60
Db 121 CGGTTTGTCATCTTCCATCCAGTACCTTGATATTTGGAAATGTATAACAGGCACAG 180
Qy 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAnLys 80
Db 181 GCCTTCCTTCTGGCAGCAGAGAGGTGCGACTTATCAAGAGGATCTCCCTCAGTGGAAACAAG 240
Qy 81 LeuLysAlaAspGluLysTrpPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 241 CTTAAAGCAGATGAGAAGTACTTCTCTCACATCTTAGCTTTTTCAGGCCAGTGAT 300
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTCCAGAGGCT 360
Qy 121 ArgCysPheTrpGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTrpSerLeu 140
Db 361 CGCTGTTTCTATGGCTTTCAAAATTCATCAGAAATGTTCAGTACAGATGTACAGTTTG 420
Qy 141 LeuIleAspThrTrpIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
```



```

division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1..1582
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DA011YH23"
                     /tissue_type="HeLa cells Cot 25-normalized"
                     /plasmid="pCMVSPORT_6"

ORIGIN
Alignment Scores:
Pred. No.:          8.95e-167      Length:          1582
Score:             1460.50         Matches:         277
Percent Similarity: 88.42%         Conservative:    36
Best Local Similarity: 78.25%       Mismatches:     35
Query Match:       80.20%          Indels:          6
DB:                3               Gaps:            1

US-10-698-228-1 (1-351) x CR608076 (1-1582)
QY      4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
Db      142 CCTGAGCGGACCCGCGTCTGCGCCAGCAAGACCGCGGAGGAGGATCTTCCAGGAGCCAC 201
QY      23 nGluSerGluIleLysSer-----AenGluGluProLeuLeuArgLysSe 38
Db      202 GGAGCGGAACTAAGCAGCTGCCCGCGGTGGAGGATGAGCCCTGCTGAGAGAAA 261
QY      38 rSerArgArgPheValIlePheProIleGlnTyPr-oAspIleTIpLysMetTyLysG 58
Db      262 CCCCCGCGCTTGTTCATCTTCCCCATCGATACCATGATATCTGCAGATGTATPAGAA 321
QY      58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
Db      322 GCGAGAGGCTTCTTTGGACCGCGGAGGAGTGGACCTCTCCAGAGCATTCAGCACTG 381
QY      78 pAsnLysLeuLysAlaAspGluLysTyPrPheIleSerHisIleLeuAlaPheAla 98
Db      382 GGAATCCCTGAAACCGGAGGAGATATTTATATCCCATGTTCTGGCTTTCTTTCAGC 441
QY      98 sSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db      442 AAGCATGGCATAGTAAATGAAACTTGTGTGAGCGATTTAGCCAGAGAGTTCAGATTAC 501
QY      118 oGluAlaArgCysPheTyrglyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db      502 AGAAGCCGCTGTTTCTATGGCTTCCAAATTGCCATGGGAAACATACATCTTGAAATGTA 561
QY      138 rSerLeuLeuIleAspThrTyIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db      562 TAGTCTTCTTATGTACACTTACATAAAGATCCCAAGAAAGGGAATTTCTCTCAATGC 621
QY      158 aIleGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
Db      622 CATTGAACAGTCCCTTGTGTCAAGAAGAGGAGACTGGCGCTTCGCTGGATTCGGGA 681
QY      178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
Db      682 CAAAGAGGCTACCTATGTGTGAACGTGTGTAGCTTTGTGCGAGTGGAGGCAATTTCTT 741
QY      198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db      742 TTCCCGGCTCTTTTGGCTCGATATTTCTGGCTCAAGAAACAGGAGCTGATGCTGGCCCTAC 801
QY      218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db      802 ATTTTCTAATGAACCTTATTAGCAGAGATGAGGTTTACACTGTGATTTGCTTGCCTGAT 861
QY      238 tPheGlnTyPrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db      862 GTTCAACACCTGGTACACAAACCATCGGAGGAGAGAGTAAAGAGAAATAATATCAATGC 921

```

```

QY      258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db      922 TGTTCGATAGAACAGAGGTTCTCTACTGAGGCCCTTGCCTGTAAGCTCATTTGGATGAA 981
QY      278 nCysIleLeuMetLysGlnTyrlleGluPheValAlaAspArgLeuLeuValGluLeuGl 298
Db      982 TTGCACCTCTAATGAAGCAATACATTTGAGTTTGTGGCAGACAGACTTATGCTTGGAACTGGG 1041
QY      298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
Db      1042 TTTTAGCAAGGTTTTTTCAGATAGAGAACCCCATTTTGACTTTATGGAGATAATTTCTACTGA 1101
QY      318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrglnArgPheAlaValMetAl 338
Db      1102 AGGAAGACTTAATCTTTTTCAGAAAGAGAGTAGGCGAGTATCAGAGGATGGGAGTGTATC 1161
QY      338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db      1162 AAGTCCCAACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1201

RESULT 11
CR602054          1588 bp      mRNA      linear      HTC 21-JUL-2004
LOCUS             full-length cDNA clone CS0DA011YH14 of Neuroblastoma of Homo
DEFINITION        sapiens (human).
ACCESSION         CR602054
VERSION           CR602054.1 GI:50482861
KEYWORDS          HTC; CNSLIT_CDNA.
SOURCE            Homo sapiens
ORGANISM          Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE          1 (bases 1 to 1588)
AUTHORS           Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE             Full-length cDNA libraries and normalization
JOURNAL            Unpublished
REMARK            Contact : Peng Liang Email : fliang@lifetech.com URL :
                  http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
                  Paraday Avenue
                  Genoscope.
REFERENCE          2 (bases 1 to 1588)
AUTHORS           Genoscope.
TITLE             Direct Submission
JOURNAL            Submitted (20-JUL-2004) Genoscope - Centre National de Sequenace :
                  BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
                  - Web : www.genoscope.cns.fr)
COMMENT            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
                  end enriched, double-strand cDNA was digested with Not I and cloned
                  into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
                  was normalized. Library was constructed by Life Technologies, a
                  division of Invitrogen.
FEATURES           Location/Qualifiers
     source        1..1588
                   /organism="Homo sapiens"
                   /mol_type="mRNA"
                   /db_xref="taxon:9606"
                   /clone="CS0DA011YH14"
                   /tissue_type="Neuroblastoma"
                   /plasmid="pCMVSPORT_6"

ORIGIN
Alignment Scores:
Pred. No.:          9e-167      Length:          1588
Score:             1460.50         Matches:         277
Percent Similarity: 88.42%         Conservative:    36
Best Local Similarity: 78.25%       Mismatches:     35
Query Match:       80.20%          Indels:          6
DB:                3               Gaps:            1

US-10-698-228-1 (1-351) x CR602054 (1-1588)
QY      4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
Db      142 CCTGAGCGGACCCGCGTCTGCGCCAGCAAGACCGCGGAGGAGGATCTTCCAGGAGCCAC 201

```


Qy	23	nglSerGluIleIysSer-----AsnGluGluProIleuLeuArgLysSe	38
Db	202	GGAGCCGAAACTAAAGCAGCTGCCCGCGTGGAGGATGAGCGCGTCTGAGAGAAA	261
Qy	38	rSerArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetYrLysG	58
Db	262	CCCCCGCGCTTTGTCACTTCCCCATCGAGTACCATGATATCTGGCAGATGATTAAGAA	321
Qy	58	nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr	78
Db	322	GGCAGAGCGTTCCTTTTGGACCGCGAGAGGTGGACCTCTCCAAGGACATTTCAGCACTG	381
Qy	78	pAsnLysLeuLysAlaAspGluLysYrPheIleSerHisIleLeuAlaPhePheAlaL	98
Db	382	GGAAATCCCTGAAACCCGAGGAGAGATATTTTATATCCCATGTTCTGCGCTTCTTTG	441
Qy	98	aSerAspGlyIleValAsnGluAenLeuValGluArgPheSerGlnGluValGlnValPr	118
Db	442	AAGCGATGGCATAGTAATGAAACTTTGGTGGAGCGATTTAGGCCAAGAAAGTTTCAG	501
Qy	118	oGluAlaArgCysPheYrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy	138
Db	502	AGRAGCCGCTGTTTCTATGGCTTCCAAATGGCATGGAAACATACATCTCTGAAATGT	561
Qy	138	rSerLeuLeuIleAspThrTyrlleArgAspProLysLeuArgGluPheLeuPheAsnAl	158
Db	562	TAGTCTTCTTATTGACACTTACATAAAGATCCCAAAGAAAGGAAATTTCTCTTCAATGC	621
Qy	158	allcGluThrMetProTyrlleLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs	178
Db	622	CATTGNAACCATGCTTGTGTCAAGAGAAGGCGACACTGGGCGCTTGGCGTGGATGGGGA	681
Qy	178	pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh	198
Db	682	CAAGAGCGCTACCTATGGTGAACGTGTGTAGCCCTTGTGCGATGGAGGACATTTCTT	741
Qy	198	eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh	218
Db	742	TTCCGGTCTTTTGGCGTCGATATTCCTGGCTCAAGAAACGAGGACTGATGCTGGCCCTC	801
Qy	218	rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe	238
Db	802	ATTTTCTTAATGAACATTATTAGCAGATGAGGGTTTACACTGTGATTTTGTGCGCTGAT	861
Qy	238	tPheGlnTyrlleuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl	258
Db	862	GTTCAAAACACTGTGTACAAACCATCGGAGGAGAGATGAAGAAATAATTTATCAATGC	921
Qy	258	aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs	278
Db	922	TGTTCCGATAGAACAGGAGTTCCTCACTGAGGCGCTTGCCTGTGAAGCTCATTTGGATGA	981
Qy	278	nCysIleLeuMetLysGlnTyrlleGluPheValAlaAspArgLeuLeuValGluLeuGl	298
Db	982	TTGCACCTCTAATGAAGCAATACATTGAGTTGTGGCAGACAGACTTATGTGGAACTGGG	1041
Qy	298	yPheSerLysValPheGlnAlaGluAenProPheAspPheMetGluAsnIleSerLeuGl	318
Db	1042	TTTTTAGCAAGGTTTTTCAGAGTAGAAGAACCCATTTCATTTATGGAGAAATATTTTCACT	1101
Qy	318	uGlyLysThrAsnPhePheGluLysArgValSerGluTyrrGlnArgPheAlaValMetAl	338
Db	1102	AGBAAAGACTTAATCTTTTGGAGAAGAGATAGGCGAGTATCAGAGGATGGGATGATGTC	1161
Qy	338	aGluThrAspAsnValPheThrLeuAspAlaAspPhe	351
Db	1162	AAGTCCAAACAGAGAATCTTTTACCTTTGGATGCTGACTTC	1201

RESULT 12

CR621427

LOCUS

DEFINITION

full-length cDNA clone CS0DE011YL04 of Placenta of Homo sapiens

1592 bp mRNA linear HTC 21-JUL-2004

```

(human).
CR621427
VERSION CR621427.1 GI:50502234
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          1 (bases 1 to 1592)
          Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
          Full-length cDNA libraries and normalization
          Unpublished
          Contact : Feng Liang Email : fliang@lifetech.com URL :
          http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
          Faraday Avenue
          2 (bases 1 to 1592)
          Genoscope.
          Direct Submission
          Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
          BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
          - web : www.genoscope.cns.fr)
          1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
          end enriched, double-strand cDNA was digested with Not I and cloned
          into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
          was normalized. Library was constructed by Life Technologies, a
          division of Invitrogen.
FEATURES             Location/Qualifiers
     source           1..1592
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DE011YL04"
                     /tissue_type="placenta"
                     /plasmid="pCMVSPORT_6"
ORIGIN
Alignment Scores:
Pred. No.:          9.04e-167          Length:          1592
Score:              1460.50            Matches:         277
Percent Similarity: 88.42%             Conservative:    36
Best Local Similarity: 78.25%           Mismatches:     35
Query Match:        80.20%              Indels:         6
DB:                 3                  Gaps:           1

US-10-698-228-1 (1-351) x CR621427 (1-1592)

Qy      4  ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
Db      153 CCTGAGCGGGACCGCGGTCTCTGCCAGCAAGACCGCGGAGGAGGATCTTCAGAGCCAC 212
Qy      23  nGluSerGluLeuLeuSer-----AsnGluGluProLeuLeuArgLysSe 38
Db      213 GGAGCCGAAACCTAAAGCAGCTGCCCCCGGTGGAGGATAGCCCTCTGAGAGAAA 272
Qy      38  rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysG 58
Db      273 CCCCCCGCGCTTTGTTCATCTCTCCCATCGAGTACCATGATATCTGGCAGATGATAAGAA 332
Qy      58  nAlaGlnAlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
Db      333 GGCAGAGGGCTTCCTTTTGACGCGCCGAGGAGGTGGACCTCTCCAAGGACATTTCAGC 392
Qy      78  pAsnLysLeuLeuAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98
Db      393 GGAATCCCTGAAACCCGAGGAGAGATATTATTATCCCAATGTTCTGGCTTCTTTGCAGC 452
Qy      98  aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db      453 AAGCGATGCGATAGTAATAATGAAACTTGTGTGGAGCGATTATGCCAAGAGATTTCAGATTAC 512
Qy      118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db      513 AGAAGCCCGCTGTTTCTATGGCTTCCAATTCGCATGGAAAAACATACATCTCTGAATGTA 572

```

```
QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
DB 573 TAGTCTTCTTATGACCTTACATAAAGATCCCAAGAAAGGAATTTCTTTCATGCG 632
QY 158 aileGluThrMetProTyrValLysLysAlaAspTrrAlaLeuArgTrrPileAlaAs 178
DB 633 CATTGAACGATGCTTGTGTCAAGAAGGCAGACATGGGCTTGGCTGGATGGGGA 692
QY 178 pArgLysSerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePh 198
DB 693 CAAAGAGGCTACCTATGCTGAACGTTGTAGACCTTGTCTGAGTGAAGGCAATTTCTT 752
QY 198 eSerGlySerPheAlaAlaIlePheTrrPheLysLysArgGlyLeuMetProGlyLeuTh 218
DB 753 TTCCGGTCTTTTGGCTCGATATTCCTCAAGAAACAGGACATGATGCTGGCTCTAC 812
QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
DB 813 ATTTTCTAATGAACCTATTAGCAGAGATGAGGTTTACACTGTGATTTTGTCTGCTGAT 872
QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluLeuIleValAspAl 258
DB 873 GTTCAACACCTGTTACACAAACCTATCGGAGGAGAGTAAAGAAATAATTTATCAATGC 932
QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
DB 933 TGTTCGGATAGAACAAGGAGTTCTCTCACTGAGGCTTGGCTGTGAAGCTCATTTGGGATGAA 992
QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG1 298
DB 993 TTGCACTCTAATGAACATAATGATTTGTGTGGCAGACACTATGCTGGAATGGG 1052
QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
DB 1053 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCATTTGACATTTATGAGAAATATTTCACTGGA 1112
QY 318 uGlyLysThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
DB 1113 AGGAAAGACTAATCTTCTTTGAGAAGAGAGTAGGCGAGTATCAGAGATGGGAGTGATGTC 1172
QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1173 AAGTCCACACAGAAATTTCTTTACCTTGGATGCTGACTTC 1212

RESULT 13
CR604378 1600 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DI037Y003 of Placenta Cot 25-normalized
DEFINITION of Homo sapiens (human).
ACCESSION CR604378
VERSION CR604378.1 GI:50485185
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1600)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1600)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
```

was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

FEATURES

source Location/Qualifiers
1..1600
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI037Y003"
/issue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Alignment Scores:

Pred. No.: 9,11e-167 Length: 1600
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 3 Gaps: 1

US-10-698-228-1 (1-351) x CR604378 (1-1600)

```
QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
DB 155 CCTGAGCGGACCCCGGTCCTGGCCAGCAAGACCGCGAGGAGGATCTTCCAGGAGCCAC 214
QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysE 38
DB 215 GGAGCCGAAACTAAAGACAGCTCCCGCCGCGTGGAGGATGAGCGCTCTGAGAGAAAA 274
QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrrPlyMetTyrLysG1 58
DB 275 CCCCCCGCGCTTGTCTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 334
QY 58 nAlaGlnAlaSerPheTrrPthrAlaGluGluValAlaAspLeuSerLysAspLeuProHisTyr 78
DB 335 GGCAGAGGCTTCTTTTGGACCCCGCGAGGAGGTGGACCTCTCCAAGGACATTCAGCACTG 394
QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
DB 395 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCATGTTCTTGGCTTTCTTTGACG 454
QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
DB 455 AAGCGATGGCATAGTAAATGAAACCTTGTGGAGCGGATTTAGCCAGAGATTCAGATTAC 514
QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
DB 515 AGAAGCCCGCTGTTCTATGCTTCCAAATTTGCCATGGAACATACATTTCTGAAATGTA 574
QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
DB 575 TAGTCTTCTTATTGACACTTACATAAAGATCCCAAGAAAGGAATTTCTTCTCAATGC 634
QY 158 alileGluThrMetProTyrValLysLysAlaAspTrrAlaLeuArgTrrPileAlaAs 178
DB 635 CATTGAACGATGCTTGTGTCAAGAAAGAGGACACTGGGCCCTTGCCTGATGGGGA 694
QY 178 pArgLysSerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePh 198
DB 695 CAAAGAGGCTACCTATGCTGAACGTTGTGTAGCCTTTTGTGCTGAGTGAAGCAATTTCTT 754
QY 198 eSerGlySerPheAlaAlaIlePheTrrPheLysLysArgGlyLeuMetProGlyLeuTh 218
DB 755 TTCCGGTCTTTTGGCTCGATATTTCTGCTCAAGAAACGAGGACTGATGCTCGCTCCAC 814
QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
DB 815 ATTTTCTAATGAACCTATTATGACAGAGATGAGGGTTTACACTGTGATTTTGTGCTGAT 874
QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
DB 875 GTTCAACACCTGTTACACAAACCATCGGAGGAGAGTAGTAAAGAAATAATTTATCAATGC 934
```

```

QY      258 aVallyleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db      935 TGTTGGATAGAACAGAGTTCTCTACTGAGGCTTGCCCTGTGAAGCTCATTTGGATGAA 994
QY      278 nCysIleLeuMetLysGlnTyriIleGluPheValAlaAspArgLeuLeuValGluLeuG1 298
Db      995 TTGCACCTAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACCTGGG 1054
QY      298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
Db      1055 TTTTACAGAGTTTTCAGAGTAGAACCACTTTGACTTTATGCGAATATTTTCACTGGA 1114
QY      318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyriGlnArgPheAlaValMetAl 338
Db      1115 AGGAAGACTAATCTTTTGAAGAGAGTACGCGAGTAGTACAGAGATGGAGTAGTATGTC 1174
QY      338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db      1175 AAGTCCACACAGAGAATCTTTTACCTTGGATGCTGACTTC 1214

RESULT 14
CR590959
LOCUS   full-length cDNA clone CS0DR013YD22 of Placenta of Homo sapiens
DEFINITION
ACCESSION   CR590959
VERSION     1 GI:50471766
KEYWORDS    HTC; CNSLT_cDNA.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 1605)
AUTHORS    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE       Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL    Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL    Full-length cDNA libraries and normalization
JOURNAL    Unpublished
REMARK     Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
            Faraday Avenue
            Genoscope.
REFERENCE   2 (bases 1 to 1605)
AUTHORS    Direct Submission
TITLE       Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL    BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES
source
    1..1605
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="CS0DE013YD22"
     /tissue_type="Placenta"
     /plasmid="pCMVSPORT_6"
ORIGIN
Alignment Scores:
Pred. No.:      9,15e-167      Length:      1605
Score:          1460.50      Matches:    277
Percent Similarity: 88.42%      Conservative: 36
Best Local Similarity: 78.25%      Mismatches: 35
Query Match:      80.20%      Indels:     6
DB:              3          Gaps:          1

US-10-698-228-1 (1-351) x CR590959 (1-1605)

QY      4 ProGluArgProGluAlaGlyLeuAspGlnAspGlu-ArgSerSerAspThrAs 23

```

```

Db      150 CTTGAGCGGACCCCGCTCTGGCCAGCAAGCCGAGGAGGATCTTTCCAGGAGCCAC 209
QY      23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db      210 GGAAGCCGAACTAAGACGAGCTGCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAAA 269
QY      38 rSerArgArgPheValIlePheProIleGlnTyriProAspIleTrpLysMetTyriLysG1 58
Db      270 CCCCAGCGCTTTTGTCTCATCTTCCCATCGAGTACCATGATATCTGCAGATGCTATAAGAA 329
QY      58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
Db      330 GGCAGAGCTTCTTTTGGACCGCGAGGAGGTGGACCTCTCCAAGGACATTCAGCACTG 389
QY      78 pAsnLysLeuLysAlaAspGluLysTyriPheIleSerHisIleLeuAlaPhePheAlaAl 98
Db      390 GGAATCCCTGAACCCGAGGAGAGATATTTATATATCCCATGTTCTTGGCTTTCTTTCGACG 449
QY      98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db      450 AAGCGATGGCATAGTAATGAAACTTGTGTGGAGCGATTAGCCAGAGATTCAGATTAC 509
QY      118 oGluAlaArgCysPheTyriGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db      510 AGAAGCCGCTGTTCTTATGGCTTCCAAATTTGCCATGGAAAACATACATCTCTGAAATGTA 569
QY      138 rSerLeuLeuLeuAspThrTyriLeArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db      570 TAGTCTCTTTATTGACACTTACATAAAGATCCCAAGAGAAAGGAATTTCTTCAATGC 629
QY      158 aIleGluThrMetProTyriValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
Db      630 CATTGAACCATGCTTGTGTCAAGAGAGAGAGAGACTGGGCTTTCCTGATTTGGGA 689
QY      178 pArgLysSerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePh 198
Db      690 CAAAGAGGCTACCTTACCTGTGTAACGCTGTTGTAGCCTTTGCTCAGTGGAGGCAATTTCTT 749
QY      198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db      750 TTCCGGTTCTTTTTCGCTCGATATTTCTGGCTCAAGAAACAGAGACTGATGCTTGGCTCAC 809
QY      218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db      810 ATTTTCTAATGAACCTTATTACGAGATGAGGGTTTACACTGTGATTTGCTTGCCTGAT 869
QY      238 tPheGlnTyriLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db      870 GTTCAACACCTGGTACACAAACCATCGAGGAGAGAGATAGAGAAATAATATCAATGC 929
QY      258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db      930 TGTTCGATAGAACAGAGATTCTCTCACTGAGGCTTGCCTGTGAAGCTCATTTGGATGAA 989
QY      278 nCysIleLeuMetLysGlnTyriIleGluPheValAlaAspArgLeuLeuValGluLeuG1 298
Db      990 TTGCACCTAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACCTGG 1049
QY      298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
Db      1050 TTTTACAGAGTTTTCAGAGTAGAACCACTTTGACTTTATGAGAGATATTTTCACTGGA 1109
QY      318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyriGlnArgPheAlaValMetAl 338
Db      1110 AGGAAGACTAATCTTTTGAAGAGAGATGAGCGAGTATCAGAGGATGGAGTAGTATGTC 1169
QY      338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db      1170 AAGTCCACACAGAGAATCTTTTACCTTGGATGCTGACTTC 1209

RESULT 15
CR614990
LOCUS   CR614990

```

QY 4 ProGluArgProGluAlaGlyLeuAspGlnAspGlu-ArgSerSerAspThrAs 23

full-length cDNA clone CS0DA009YB23 of Neuroblastoma of Homo sapiens (human).
 DEFINITION
 CR614990
 VERSION
 CR614990.1 GI:50495797
 KEYWORDS
 HTC; CNSLT_cDNA.
 SOURCE
 Homo sapiens
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 1612)
 AUTHORS
 Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
 TITLE
 Full-length cDNA libraries and normalization
 JOURNAL
 Unpublished
 CONTACT
 Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue
 REFERENCE
 2 (bases 1 to 1612)
 AUTHORS
 Genoscope.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - web : www.genoscope.cns.fr)
 COMMENT
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
 FEATURES
 Location/Qualifiers
 1..1612
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DA009YB23"
 /tissue_type="Neuroblastoma"
 /plasmid="pCMVSPORT_6"
 ORIGIN
 Alignment Scores:
 Pred. No.: 9,21e-167 Length: 1612
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 3 Gaps: 1
 US-10-698-228-1 (1-351) x CR614990 (1-1612)
 QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
 DB 152 CCTGAGCGGGACCCGGCTCTGGCCAGCAAGACCGCGAGGAGATCTTCCAGGAGCCAC 211
 QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
 DB 212 GGAGCCGAAACCTAAAGACGCTGCCCGCGGTGGAGGATGAGCCGCTGTGAGAGAAA 271
 QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysG1 58
 DB 272 CCCCCCGCGCTTGTTCATCTCCCATCGAGTACCATGATATCTGCAGATGTATAAGAA 331
 QY 58 nAlaGlnAlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
 DB 332 GCGAGAGGCTTCTTTTGGACCGCGAGGAGGTGGACCTCTCCAGAGACATTCAGGACTG 391
 QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
 DB 392 GGAATCCCTGAAACCCGAGAGAGATATTTATATATCCATGTTCTGCTTTCTTCGAGC 451
 QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
 DB 452 AAGCGATGCATAGTAATAATGAAACCTTGGTGGAGCGGATTTAGCCAAAGAGTTCAGATTAC 511
 QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138

512 AGAAGCCCGCTGTTTCTATGGCTTCCAAATGGCCATAAAGAAACATACATCTTCTGAATGTA 571
 QY 138 rSerLeuLeuLeuAspThrTyrIleArgAspProLysLysValArgGluPheLeuPheAsnAl 158
 DB 572 TAGTCTTCTTATTGACACITACATAAAGATCCCAAGAAAGGGAATTTCTTCCATGTC 631
 QY 158 aIleGluThrMetProTyrValLysLysLysAlaAspTirAlaLeuArgTirPilleAlaAs 178
 DB 632 CATTTGAAACGATGCTTGTCTCAAGAAAGAGGAGAGCTGGGCTTGGCTGGATTGGGA 691
 QY 178 pArgLysSerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePh 198
 DB 692 CAAAGAGGCTACCTATGCTGAACGCTGTGTAGCCCTTGTCTGCTGAGGAGCAATTTCTT 751
 QY 198 eSerGlySerPheAlaAlaIlePheTyrPheLysLysArgGlyLeuMetProGlyLeuTh 218
 DB 752 TTCGGTCTTCTTTTGGCTGATATTTCTGGCTCAAGAAACGAGAGCTGATGCTTGGCTCAC 811
 QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 DB 812 ATTTTCTAATGAATATTATAGCAGAGATGAGGTTTACACTGTGATTTGCTTGCCTGAT 871
 QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
 DB 872 GTTCAAAACACCTGCTGACACAAACCATCGAGGAGAGATAGAGAAATAATATATCAATGC 931
 QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 DB 932 TGTTGGATAGAACAGAGATTCCTCACTGAGGCTTGCCTGTGAGGCTCATTTGGATGAA 991
 QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluGluG1 298
 DB 992 TTGCACCTTAATCGACCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAATGGG 1051
 QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
 DB 1052 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCATTTGACTTTTATCGAGAAATATTTCACTGA 1111
 QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
 DB 1112 AGGAAAGACTTAACCTTCTTTCAGAAAGAGAGATGAGGCGAGTATCAGAGGATGGGAGTATGTC 1171
 QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1172 AAGTCCAAACAGAGAAATTTCTTTTACCTTGGATGCTGACTTC 1211

Search completed: October 30, 2005, 08:53:08
 Job time : 2604 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 30, 2005, 07:02:47 ; Search time 156 Seconds
(without alignments)

3681.623 Million cell updates/sec

Title: US-10-698-228-1

Perfect score: 1821

Sequence: 1 MGDPERPEAAGLDQDERSS.....QRFVMAETTDNVFTLDADF 351

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO spool/US10698228/runat_26102005_100715_5137/app_query.fasta_1.519
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228 @cgn_1_1_54 @runat_26102005_100715_5137 -NCPUS=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*

1: /cgn2_6/prodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/prodata/1/ina/PCUTUS_COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	100.0	1053	4	US-10-019-733-2
2	1821	100.0	1053	4	US-10-019-733-12
3	1821	100.0	1081	4	US-10-019-733-4
4	1821	100.0	4955	4	US-10-019-733-3
5	1460.5	80.2	2479	4	US-09-949-016-2025
6	1460.5	80.2	2500	4	US-09-962-665-9
7	1460.5	80.2	2500	4	US-09-023-655-1370
8	1460.5	80.2	2500	4	US-09-963-333-9
9	1460.5	80.2	2500	4	US-09-962-677-9
10	1460.5	80.2	2500	4	US-09-949-016-145
11	1210	66.4	14176	1	US-08-307-499-1
12	1210	66.4	14176	1	US-08-307-499-14

C	13	1210	66.4	14176	3	US-09-299-268-1	Sequence 1, Appli
	14	1210	66.4	14176	3	US-09-299-268-14	Sequence 14, Appl
	15	1131	62.1	1236	4	US-09-248-796A-3914	Sequence 3914, Ap
	16	1072.5	58.9	825	1	US-08-307-499-40	Sequence 40, Appl
	17	1072.5	58.9	825	3	US-09-299-268-40	Sequence 40, Appl
	18	984.5	54.1	1112	1	US-08-136-743B-1	Sequence 1, Appli
	19	969.5	53.2	1083	4	US-09-248-796A-3915	Sequence 3915, Ap
	20	768	42.2	481	3	US-08-905-223-125	Sequence 125, App
	21	674	37.0	656	3	US-09-328-111-791	Sequence 791, App
	22	599	32.9	449	4	US-09-270-767-14670	Sequence 14670, A
	23	558	30.6	419	3	US-08-905-223-129	Sequence 129, App
	24	543.5	29.8	11820	4	US-09-949-016-11887	Sequence 11887, A
	25	543.5	29.8	11826	4	US-09-949-016-13767	Sequence 13767, A
C	26	513.5	28.2	13706	4	US-09-902-540-1124	Sequence 1124, Ap
	27	509.5	28.0	1053	4	US-09-902-540-2666	Sequence 2666, Ap
C	28	488.5	26.8	366	1	US-08-307-499-13	Sequence 13, Appl
	29	488.5	26.8	366	3	US-09-299-268-13	Sequence 13, Appl
C	30	392.5	21.6	297	4	US-09-313-294A-4435	Sequence 4435, Ap
	31	381.5	21.0	1338	4	US-09-328-352-588	Sequence 588, App
C	32	378.5	20.8	1707	4	US-09-252-991A-3180	Sequence 3180, Ap
	33	377.5	20.7	1431	4	US-09-252-991A-2787	Sequence 2787, Ap
C	34	370	20.3	984	4	US-09-252-991A-2692	Sequence 2692, Ap
	35	355	19.5	373	4	US-09-401-064-319	Sequence 319, App
C	36	355	19.5	373	4	US-09-401-064-337	Sequence 337, App
	37	338	18.6	3278	1	US-07-778-890A-1	Sequence 1, Appli
	38	335	18.4	1230025	4	US-09-198-452A-1	Sequence 1, Appli
	39	335	18.4	1230230	4	US-09-438-185A-1	Sequence 76, Appli
C	40	330	18.1	124884	4	US-09-661-596A-76	Sequence 76, Appli
C	41	330	18.1	124884	4	US-09-913-514-1	Sequence 1, Appli
C	42	330	18.1	125157	4	US-09-913-514-2	Sequence 2, Appli
	43	317	17.4	3678	4	US-09-112-580-13	Sequence 13, Appl
C	44	307	16.9	35100	2	US-08-770-379-19	Sequence 19, Appl
C	45	307	16.9	35100	3	US-08-757-669A-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1

US-10-019-733-2

; Sequence 2, Application US/10019733

; Patent No. 6682917

; GENERAL INFORMATION:

; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA

; TITLE OF INVENTION: New Protein and its DNA

; FILE REFERENCE: 2619W00P

; CURRENT APPLICATION NUMBER: US/10/019,733

; PRIOR FILING DATE: 2001-12-28

; PRIOR APPLICATION NUMBER: JP 11-181131

; PRIOR FILING DATE: 1999-06-28

; PRIOR APPLICATION NUMBER: JP 11-192391

; PRIOR FILING DATE: 1999-07-06

; PRIOR APPLICATION NUMBER: JP 2000-017770

; PRIOR FILING DATE: 2000-01-21

; NUMBER OF SEQ ID NOS: 14

; SEQ ID NO 2

; LENGTH: 1053

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION:

US-10-019-733-2

Alignment Scores:

Pred. No.: 6.06e-227

Score: 1821.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 100.00%

DB: 4

US-10-698-228-1 (1-351) x US-10-019-733-2 (1-1053)

QY

1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20

QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGTCTATATTCTGGCTAAAGAGAGAGGTCCTTATGCCAGGACTCACATTTTTC 660
QY 221 AsnGluLeuLysSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleLeuValAspAlaValLys 260
Db 721 TACTTAGTAAATAGCCTTCAGAGAAAGGTCAGGAGATCATGTTGATGCTGCAAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 781 ATTGAGCAGGAGTCTTTTACAGAGCCTTGCCAGTTGGCTCATTTGGAATGAATTCATT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 841 TTGATGAACACAGTACATAGTGTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 901 AAGGTTTTTTCAGCGAGAAATCCTTTTGATTTTATGGAACAATTTCTTTAGAAAGAAA 960
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 961 ACAAAATTTCTTGAGAAACGAGTTTCAGAGATATCAGCGTTTGCATTTATGCGAGAAACC 1020
QY 341 ThrAsnValPheThrLeuAspAlaAspPhe 351
Db 1021 ACAGATAACGCTCTCACCTTGGATGCGAGATTTT 1053

RESULT 3

US-10-019-733-4
; Sequence 4, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-4

Alignment Scores:

Pred. No.: 6,32e-227 Length: 1081
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-698-228-1 (1-351) x US-10-019-733-4 (1-1081)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 20 ATGGCGCGACCCGGAAGCGCGGATCAGGATCAGGATCATCTTCA 79
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40

Db 80 GACACCAACGAAAGTGAATAAAGTCAAATCAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 139
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Db 140 CGGTTTGTCTATCTTCCAAATCCAGTACCCTGATATTTGGAAAAATGATAAAGACGACAG 199
QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db 200 GCTTCTCTCTGGACAGCAGAGAGGTTGACTTATCAAAGGATCTCCCTCACTGGAACAAG 259
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAspSerAsp 100
Db 260 CTTAAAGCAGATCAGAAAGTACTTCTCATCTCACATCTTAGCCTTTTTCAGCCAGCATGAT 319
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 320 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGTTCCAGAGGCT 379
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 380 CGCTGTTTCTATGCTTCCAAATTCATCGAGATGTTCACTCAGAGATGTCAGATTG 439
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGCGAAATTTTATTAATGCAATTGAA 499
QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 500 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCCCTTGGATGGATGAGTAGAAAA 559
QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db 560 TCTACTTTTGGGAAAGAGTGGTGGCTTCTCTCTGTAGAGAGATTCTTCTCAGA 619
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 620 TCTTTTGTCTATATTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACATTTTTC 679
QY 221 AsnGluLeuLysSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 680 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 739
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 740 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA 799
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 800 ATTGAGCAGGAGTCTTTTACAGAAAGCCTTGGCAGTTGGCCTCATTTGGAATGAATTGCAT 859
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 860 TTGATGAAAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGGAACCTGGATTTCTCA 919
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 920 AAGGTTTTTTCAGCGAGAAATCCTTTTGATTTTATGGAACAATTTCTTTAGAAAGAAAA 979
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 980 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGATGATGCGAGAAACC 1039
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1040 ACAGATAACGCTCTTCCACTTGGATGCGAGATTTT 1072

RESULT 4

US-10-019-733-3
; Sequence 3, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA

QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 Db ATTTTCTAATGAATGATTTATTAGCAGAGATGAGGGTTTACACTGTGATTTTCTGCTGCTGAT 1021
 QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluLeuIleValAspAl 258
 Db TTTTCAAAACACCTGGTGTACAAACCATCGGAGGAGAGAGTAAAGAGAAATAATTATCAATGC 1081
 QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 Db TGTTCGGATAGAACAGGAGTTCCTCACTGAGGCTTGCCTGTGAAGCTCAATGGGATGAA 1141
 QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuG1 298
 Db TTGCACCTCTAATGAAGCAATACATTGAGTTTGTGCGAGAGAGACTTATGCTGGAACCTGGG 1201
 QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
 Db TTTTAGCAAGGTTTTCAGAGTAGAGAACCCATTGTGACTTTATGGAGAAATATTTCACTGGA 1261
 QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
 Db AGGAAGACTAATCTTTTGTAGAGAGAGAGTAGCGGAGTATCAGAGGATGGAGTGATGTC 1321
 QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db AAGTCCAACAGAGAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 8

US-09-963-333-9
 ; Sequence 9, Application US/09963333
 ; Patent No. 6664062
 ; GENERAL INFORMATION:
 ; APPLICANT: Stanton, Jr., Vincent P.
 ; TITLE OF INVENTION: THYMIDINE SYNTHASE GENE SEQUENCE VARIANCES
 ; TITLE OF INVENTION: HAVING UTILITY IN DETERMINING THE TREATMENT
 ; TITLE OF INVENTION: OF DISEASE
 ; FILE REFERENCE: 11926-015002
 ; CURRENT APPLICATION NUMBER: US/09/963,333
 ; CURRENT FILING DATE: 2001-09-24
 ; PRIOR APPLICATION NUMBER: 09/658,659
 ; PRIOR FILING DATE: 2000-09-08
 ; PRIOR APPLICATION NUMBER: 09/596,033
 ; PRIOR FILING DATE: 2000-06-15
 ; PRIOR APPLICATION NUMBER: 09/357,743
 ; PRIOR FILING DATE: 1999-07-20
 ; PRIOR APPLICATION NUMBER: 09/357,024
 ; PRIOR FILING DATE: 1999-07-19
 ; PRIOR APPLICATION NUMBER: 60/093,484
 ; PRIOR FILING DATE: 1998-07-20
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 9
 ; LENGTH: 2500
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 128, 1464
 ; OTHER INFORMATION: n = g or a
 ; NAME/KEY: misc_feature
 ; LOCATION: 189
 ; OTHER INFORMATION: n = t or g
 ; NAME/KEY: misc_feature
 ; LOCATION: 524
 ; OTHER INFORMATION: n = c or g
 ; NAME/KEY: misc_feature
 ; LOCATION: 1399
 ; OTHER INFORMATION: n = t or a
 ; NAME/KEY: misc_feature
 ; LOCATION: 1636, 1738, 2259
 ; OTHER INFORMATION: n = c or t

Alignment Scores:
 Pred. No.: 2,09e-179 Length: 2500
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 4 Gaps: 1

US-10-698-228-1 (1-331) x US-09-963-333-9 (1-2500)

QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
 Db CCTGAGCGGACCGCGCTCTGGCGCAGCAACCGAGGAGGATCTTCCAGGAGCCAC 361
 QY 23 nGluSerGluIleLysSer-----AsnGluLupProLeuLeuArgLysSe 38
 Db GGAGCCGAAACTAAAGCAGCTGCCCGCGCGTGGAGGATGAGCGCTGCTGAGAGAAA 421
 QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysG1 58
 Db CCCCCCGCGCTTTGTCACTTCCCATCGAGTACATGATCTGGCAGATGTATAAGAA 481
 QY 58 nAlaGlnAlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
 Db GGCAGAGGCTTCTTTTGGACCGCGAGGAGTTGACCTCTCNAAGGACATTCAGCACTG 541
 QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
 Db GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCCATCTTCTGGCTTTCTTTGCAGC 601
 QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
 Db AAGCATGGCATAGTAAATGAAACTTGTGTGGAGCGATTTAGCCAAAGAGTTTCAGATTAC 661
 QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
 Db AGAAGCCCGCTGTTCTATGGCTTCCAAATGCCATGGGAAACATACATCTGAAATGTA 721
 QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
 Db TAGTCTTCTTATTGACACTTACATAAAAGATCCCAAGAGAAAGGAAATTTCTTCAATGC 781
 QY 158 alieGluThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrIleAlaAs 178
 Db CATTGAAACGATGCTTGTGTCAAGAAAGAGGAGACTGGGCGCTTGGCGTGGATGGGA 841
 QY 178 pArgLysSerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePh 198
 Db CAAAGAGGCTACCTATGTTGAACTGTTGTAGCCCTTTGCTCAGTGGAGGCAATTTCTT 901
 QY 198 eSerGlySerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 Db TTCGGGTTCTTTTGGTCGATATTTCTGGCTCAAGAAACAGAGACTGATGCTGGCTCAC 961
 QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 Db ATTTTCTAATGAACCTTATTAGCAGATGAGGGTTTACACTGTGATTTTGTGCTGCTGAT 1021
 QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluLeuIleValAspAl 258
 Db GTTCAAAACACCTGGTACAAACCATCGGAGGAGAGAGTAAAGAGAAATAATTATCAATGC 1081
 QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 Db TGTTCGGATAGAACAGGAGTTCCTCACTGAGGCTTGCCTGTGAAGCTCAATGGGATGAA 1141
 QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuG1 298
 Db TTGCACCTCTAATGAAGCAATACATTGAGTTTGTGGCAGAGAGACTTATGCTGGAACCTGG 1201
 QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
 Db AAGTCCAACAGAGAATCTTTTACCTTGGATGCTGACTTC 1361

Db 1202 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCATTGTGACTTTATGGAGATATTTCACTGGA 1261
Qy 318 uGlyLysThrAsnPhenPheGluLysArgValSerGluLysGlnArgPheAlaValMetal 338
Db 1262 AGGAAAGACTAACTCTTTTGAGAGAGAGTAGCGGAGTATCAGAGGATGGAGTGTATGC 1321
Qy 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1322 AAGTCCACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 9

US-09-962-677-9
; Sequence 9, Application US/09962677
; Patent No. 6759200
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE PHOSPHORYLASE GENE SEQUENCE
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING
; TITLE OF INVENTION: THE TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015003
; CURRENT APPLICATION NUMBER: US/09/962,677
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128, 1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc_feature
; LOCATION: 1636, 1738, 2259
; OTHER INFORMATION: n = c or t
US-09-962-677-9

Alignment Scores:
Pred. No.: 2.09e-179 Length: 2500
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 4 Gaps: 1

US-10-698-228-1 (1-351) x US-09-962-677-9 (1-2500)

Qy 4 ProGluArgProGluAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspPhras 23
Db 302 CTGACGGAGACCCGGCTCTGCGCCAGAACACCGGAGGAGGATCTTCCAGGAGCCAC 361
Qy 23 nGluSerGluLeuLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 362 GGAGCCGAAACTAAAGCAGCTGCCCGCGGTGGAGGATGAGCCGCTCTGAGAGAAA 421

Qy 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGl 58
Db 422 CCCCCCGCGTGTGTATCTTCCCATCGAGTACCATGATATCTGGCAGATGATAGA 481
Qy 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
Db 482 GGCAGAGGCTTCTTTTGGACCGCCGAGGAGGTTGACCTCTCNAAGGACATTCAGC 541
Qy 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
Db 542 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCATCTTCTGGCTTCTTTG 601
Qy 98 aSerAspGlyIleValAlaAsnGluLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 602 AGCGATGGCATAGTAATGAAACCTTGGTGGAGCGAATTAGCCAAGAGTTCAATTAC 661
Qy 118 oGluAlaArgCysPheTrpGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 662 AGAAGCCCGCTGTTTCTATGCTTCCAAATTTGCCATGGAAACATACATTTCTGAA 721
Qy 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db 722 TAGTCTTCTTATTGACACTTACATAAAAGATCCCAAGAGAAAGGAAATTTCTTCA 781
Qy 158 aIleGluThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
Db 782 CATTTAAACCATGCTTGTCTCAAGAGAGAGAGACTGGGCGCTTGCCTGATGGGA 841
Qy 178 pArgLysSerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePh 198
Db 842 CAAAGAGGCTACCTATGTTGACGCTGTTGTAGCCTTTGCTCAGTGGAGGCAATTTCT 901
Qy 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db 902 TTCGGTCTTTTTCGCTCGATATTTCTGGCTCAAGAAACGAGGACTGATGCTGGCTC 961
Qy 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db 962 ATTTTCTAATGAACCTTATTAGCAGATGAGGGTTTACACTGTGATTTTGTCTGCTG 1021
Qy 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db 1022 GTTCAACACCTGGTACACAAACCATCGGAGGAGAGTAGAGAGAAATAATTATCATGC 1081
Qy 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMeAs 278
Db 1082 TGTTCGGATAGAACAGAGGTTCTCACTGAGGCGCTTGCCTGGAAGCTCAITGGGATGA 1141
Qy 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGl 298
Db 1142 TTGCACCTTAATGAAGCAATACATTTAGTTTGTGGCAGACAGACTTATGTGGAAC 1201
Qy 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
Db 1202 TTTTAGCAAGTCTTTCAGAGTAGAGAACCCATTTGACTTTATGGAGATATTTCACTG 1261
Qy 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMeAl 338
Db 1262 AGGAAAGACTAACTCTTTTTCAGAGAGAGTAGGCGAGTATCAGAGGATGGGAGTGC 1321
Qy 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1322 AAGTCCACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 10

US-09-949-016-145
; Sequence 145, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 145
LENGTH: 2500
TYPE: DNA
ORGANISM: Human
US-09-949-016-145

Alignment Scores:
Pred. No.: 2,096-179 Length: 2500
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 4 Gaps: 1

US-10-698-228-1 (1-351) x US-09-949-016-145 (1-2500)

QY 4 ProGluArgProGluAlaGlyLeuAspGluArgSerSerAspThrAs 23
DB 302 CCGAGCGGAGCCCGCGTCTCCGCCAGCAGACCGCGGAGGAGTCTCCAGGAGCCAC 361
QY 23 nGluSerGluLeuLysSer-----AsnGluGluProLeuLeuLys 38
DB 362 GGAGCGGAAACTAAGCAGAGTGGCCCGCGGTGGAGGATGAGCGCTCTCAGAGAAA 421
QY 38 rSerArgArgPheValIlePheProIleGlnTyProAspIleTrpLysMetTyLysG 58
DB 422 CCCCAGCGCTTGTCACTTCCCATCGATCCATGATATCTGCAGATGTATAGAA 481
QY 58 nAlaGlnAlaSerPheTrpThrAlaGluValAlaAspLeuSerLysAspLeuProHisTr 78
DB 482 GGCAGAGGCTTCTTTGGACCGCGGAGGAGGTGACCTCTCCAGGACATTCAGCAC 541
QY 78 pAsnLysLeuLysAlaAspGluLysTyPheIleSerHisIleLeuAlaPhePheAla 98
DB 542 GGAATCCCTGAAACCGAGGAGATATTTATATCCCATGTCTGGCTTTCTTTCAGC 601
QY 98 aSerAspGlyIleValAsnGluAlaValGluArgPheSerGlnGluValGlnValPr 118
DB 602 AAGCATGGCATAGTAAATGAAACCTTGGTGGAGCGATTTAGCAGAAAGTTCAGATTAC 661
QY 118 oGluAlaArgCysPheTyGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
DB 662 AGAAGCCCGCTCTTTCTATGCTTCCAAATTCCTGGAACATACATCTCTGAAATGTA 721
QY 138 rSerLeuLeuIleAspThrTyIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
DB 722 TAGTCTTCTTATGTACACTTACATAAAGATCCCAAGAAAGGAATTTCTCTTCAATGC 781
QY 158 aIleGluThrMetProTyValIleLysLysValaAspTrpAlaLeuArgTrpIleAla 178
DB 782 CATTAAGACATGCTTGTGTCAAGAAAGGACAGCTGGGCTTGGCGTGGATGGGGA 841
QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaValAluGluGlyValPhePh 198
DB 842 CAAAGAGGCTACCTATGGTGAACGTTGTAGSCCTTTGCTGAGTGAAGGACATTTCTT 901
QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
DB 902 TTCGGTCTTCTTGGTGCATATCTTGGCTCAAGAAACGAGGACTGATCGCTGGCCTCAC 961
QY 218 rPheSerAnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238

DB 962 ATTTTCTAATCAACTTATTAGCAGAGATGAGGTTTACACTGTGATTTTGTCTGCTGAT 1021
QY 238 tPheGlnTyLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
DB 1022 GTTCAAAACCTGTGTACAAACCATCGAGGAGAGATAGAGAAATAATATCAATGC 1081
QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
DB 1082 TGTTCGATAGAACAGGAGTCTCTCACTGAGGCTTGCCTGTGAAGCTCAITGGATGAA 1141
QY 278 nCysIleLeuMetLysGlnTyIleGluPheValAlaAspArgLeuValGluLeuGl 298
DB 1142 TTGCACCTTAATGAAGCAATACATTTGATTTTGTGCACAGACACTTATCTGGAAC 1201
QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAnIleSerLeuGl 318
DB 1202 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCCTTTGACTTTTATGGAGATAATTTCTAC 1261
QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyGlnArgPheAlaValMetAl 338
DB 1262 AGGAAGACTTACTTCTTTGAGAAGACAGATAGGCGGATATCAGAGGATGGGATGATGTC 1321
QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1322 AAGTCCAACAGAGAAATTTTACCTTGGATGCTGACTTC 1361

RESULT 11
US-08-307-499-1/c
Sequence 1, Application US/08307499
Patent No. 5651972
GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Vi uela, Eladio
APPLICANT: Gibbs, E.P.J.
TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
TITLE OF INVENTION: Live Vaccine Vector
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/307,499
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,241
FILING DATE: 1-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,630
FILING DATE: 29-JUN-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/342,212
FILING DATE: 21-APR-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

LENGTH: 14176 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:

NAME/KEY: CDS

LOCATION: 3852..4226

FEATURE:

NAME/KEY: CDS

LOCATION: 4585..4887

FEATURE:

NAME/KEY: CDS

LOCATION: 5131..5310

FEATURE:

NAME/KEY: CDS

LOCATION: 5760..5912

FEATURE:

NAME/KEY: CDS

LOCATION: 6786..7130

FEATURE:

NAME/KEY: CDS

LOCATION: 10148..10513

US-08-307-499-1

Alignment Scores:

Pred. No.: 1,37e-145 Length: 14176
Score: 1210.00 Matches: 233
Percent Similarity: 82.01% Conservative: 36
Best Local Similarity: 71.04% Mismatches: 58
Query Match: 66.45% Indels: 2
DB: 1 Gaps: 1

US-10-698-228-1 (1-351) x US-08-307-499-1 (1-14176)

QY 24 GluSerGluIleuSerAsnGluGluProLeuLeuArgLysSerSerArgArgPheVal 43
DB 10842 AAAAATGATTTTATACAGATGAGCTATCTTCAAGAGCTGATCTAGTTCGTT 10783
QY 44 IlePheProIleGlnTyrProIleTyrLysMetTyrLysGlnAlaGlnAlaSerPhe 63
DB 10782 ATTTCCCTATTAAAGTATCATGATATCTGAAAATGTATAAACAATCAGTGGCAAGTTT 10723
QY 64 ThrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLysLeuLysAla 83
DB 10722 TGACCGTTGAGAGAGTAGATTTATCAAAAGATTTAGATGATTGGGATAAATTAACATAA 10663
QY 84 AspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAspGlyIleVal 103
DB 10662 GACGAAAATACTTTATAAACAATATAGCATTTTTCATCTAGTGTGATTTGTA 10603
QY 104 AsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAlaArgCysPhe 123
DB 10602 AATGAGATTTAGCGGAAAGATTTTATGTGGAGTACAGTGTTCAGAGCCAGCATGTTTC 10543
QY 124 TyrGlyPheGlnIleLeuLeuGluAsnValHisSerGluMetTyrSerLeuLeuIleAsp 143
DB 10542 TATGGATTTCAATAGCTATGGAATAATATCATTCAGAAATGTATAGTTTATTAATAGAT 10483
QY 144 ThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGluThrMetPro 163
DB 10482 ACNATGTAGAGATATATAGAAAATATGCATTTTATTAACCTATAGAAACAATGGAA 10423
QY 164 TyrValLysLysAlaAspTrrpAlaLeuArgTrrpIleAlaAspArgLysSerThrPhe 183
DB 10422 TCGTAAAAAAGAAAGCTGATTGGGCCAGAAATGGATATCTAGCAACAG--GTATAT 10366
QY 184 GlyGluArgValAlaPheAlaAlaValGluGluValPhePheSerGlySerPheAla 203
DB 10365 GGAGAAAGAGTAGTAGCATTTGACAGCTGTGGAGGAATATCTTTCTGTTTCATTGCT 10306
QY 204 AlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeu 223

DB 10305 GCTATATTTTGGATAAAAAACGAGGATTCATCCCGGATTAACATTTTCTTAATGAACATA 10246
QY 224 IleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGlnTyrLeuVal 243
DB 10245 ATAAGTAGAGACGAAAGGTTTACATTTGATTTTCGGTGTAAATGTTTAAACATTTATTA 10186
QY 244 AsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLysIleGluGln 263
DB 10185 CATCCACCATCTAAGGAAGTTATAACGTCGATATCATGTCGGTTAATATAGAAAAG 10126
QY 264 GluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIleLeuMetLys 283
DB 10125 GAGTTTTTGACAGTTGCTATTCGGTGCATCTTATAGTATGAATGTTGTTTAAATGCT 10066
QY 284 GlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSerLysValPhe 303
DB 10065 CAGTATATAGAAATTCGTCGAGATAGATTATTAACAGAGTTAGTTGTGAAAAGT--CTC 10008
QY 304 GlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLysThrAsnPhe 323
DB 10007 AATGTATATATCTTTTACGTTTATGAGTATATATACAGAGTTAGTTGTGAAAAGT--CTC 9948
QY 324 PheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThrThrAspAsn 343
DB 9947 TTCGAACGACGAGTTAGTGAATATCAAAAGATGGGGTGTTCACAAATAAAGAGAGAAT 9888
QY 344 ValPheThrLeuAspAlaAspPhe 351
DB 9887 ATATTTCTACGGATATAGATTTT 9864

RESULT 12

US-08-307-499-14

; Sequence 14, Application US/08307499

; Patent No. 5651972

; GENERAL INFORMATION:

; APPLICANT: Moyer, Richard W.

; APPLICANT: Vi uela, Eladio

; APPLICANT: Gibbs, E.P.J.

; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a

; NUMBER OF SEQUENCES: 60

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: David R. Saliwanchik

; STREET: 2421 N.W. 41st Street, Suite A-1

; CITY: Gainesville

; STATE: Florida

; COUNTRY: U.S.A.

; ZIP: 32606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/307,499

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/908,241

; FILING DATE: 1-JUL-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/908,630

; FILING DATE: 29-JUN-1992

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/342,212

; FILING DATE: 21-APR-1992

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Saliwanchik, David R.

; REGISTRATION NUMBER: 31,794

; REFERENCE/DOCKET NUMBER: UF35.1.FWCCI

; TELECOMMUNICATION INFORMATION:

Patent No. 6217882
 GENERAL INFORMATION:
 APPLICANT: Moyer, Richard W.
 APPLICANT: Viuela, Eladio
 APPLICANT: Gibbs, E.P.J.
 TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
 TITLE OF INVENTION: Live Vaccine Vector
 NUMBER OF SEQUENCES: 60
 CORRESPONDENCE ADDRESS:
 ADDRESSER: David R. Saliwanchik
 STREET: 2421 N.W. 41st Street, Suite A-1
 CITY: Gainesville
 STATE: Florida
 COUNTRY: U.S.A.
 ZIP: 32606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/299,268
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/901,127
 FILING DATE:
 APPLICATION NUMBER: US 07/908,241
 FILING DATE: 1-JUL-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/908,630
 FILING DATE: 29-JUN-1992
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/342,212
 FILING DATE: 21-APR-1992
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Saliwanchik, David R.
 REGISTRATION NUMBER: 31,794
 REFERENCES/DOCKET NUMBER: UF35.1.FWCC1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 904-375-8100
 TELEFAX: 904-372-5800
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 14176 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 3852..4226
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 4585..4887
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 5131..5310
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 5760..5912
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 6786..7130
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 10148..10513
 US-09-299-268-1

Alignment Scores:

Pred. No.: 1.37e-145 Length: 14176

Score: 1210.00 Matches: 233
 Percent Similarity: 82.01% Conservative: 36
 Best Local Similarity: 71.04% Mismatches: 58
 Query Match: 66.45% Indels: 2
 DB: 3 Gaps: 1
 US-10-698-228-1 (1-351) x US-09-299-268-1 (1-14176)
 QY 24 GluSerGluLeuLysSerAsnGluProLeuLeuArgLysSerSerArgPheVal 43
 DB 10842 AAAATGATTTTATACGAATGGAGCCCTATCTTCAAGAGCTGATTTCTAGGTTCGTT 10783
 QY 44 IlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGlnAlaSerPhe 63
 DB 10782 ATTTCCCTATTAAAGTATCATGATATCTGGAATGTATAACAATCAGTGGCAAGTTT 10723
 QY 64 TrpThrAlaGluLeuValAspLeuSerLysAspLeuProHisTrpAsnLysLeuLysAla 83
 DB 10722 TGGACCGTTGAAGAAGTAGATTTATCAAAAGATTAGATGATTTGGGATAAATTAACATAA 10663
 QY 84 AspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAspGlyIleVal 103
 DB 10662 GACGAAAAATACTTTATAAACATATACTAGCATTTTGCATCTAGTATGGTATTGTA 10603
 QY 104 AsnGluAsnLeuValGluArgPheSerGlnGlnValGlnValProGluAlaAlaArgCysPhe 123
 DB 10602 AATGACAATTTAGCGGAAGATTTTATGTGATGTACAGTGTTCAGAGGCACGATGTTTC 10543
 QY 124 TyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeuLeuIleAsp 143
 DB 10542 TATGGATTTCAAAATAGCTATGGAAAAATATTCATTCAGAAATGTATAGTATTATTAATAGAT 10483
 QY 144 ThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGluThrMetPro 163
 DB 10482 ACATATGTAAAGATAATATAGAAAAAATGCAATTTATTTAACGCTATAGAAACAAATGAA 10423
 QY 164 TyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLysSerThrPhe 183
 DB 10422 TCGTAAAAAAGAAAGCTGATTTGGCCAGAAATGGATATCTAGCAACAAG--GTATAT 10366
 QY 184 GlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGlySerPheAla 203
 DB 10365 CGAGAAAGAGTAGTAGCATTTGCAGCTGTGAGGGAATATCTTTCTGTTCAATTTGCT 10306
 QY 204 AlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeu 223
 DB 10305 GCTATATTTTGGATAAAAAAACGAGGATGATGCCCGGATTAACATTTCTTAATGAACATA 10246
 QY 224 IleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGlnTyrLeuVal 243
 DB 10245 ATATGTAGAGACGAGGTTTACATTTGTGATTTTGGGTGTTTAAATGTTTAAACATTTATTA 10186
 QY 244 AsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLysIleGluGln 263
 DB 10185 CATCCACCCTCTAAGGAAGTTATAACGTCGATATCATTCGTCGGTAAATATATAGAAAAG 10126
 QY 264 GluPheLeuThrGluAlaLeuProValGlyIleGlyMetAsnCysIleLeuMetLys 283
 DB 10125 GAGTTTTTGACAGTTGCTATTCGCGTGATCTTATAGGTATGAATTTGTTTAAATGCT 10066
 QY 284 GlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSerLysValPhe 303
 DB 10065 CAGTATATAGAAATTCGTCGAGATAGATTTAATACAGAGTTAGTTGTTGGAAGAGT--CTC 10008
 QY 304 GlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLysThrAsnPhe 323
 DB 10007 AATGTATATAATCCTTTTACGTTTATCGAGTATATATCACTAGAAAGGTAAAGACTAATTTT 9948
 QY 324 PheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThrThrAspAsn 343
 DB 9947 TTCGAACGACGAGTTAGTGAATATCAAAAGATGGGGGTGTTTACAAAATAAAGAGAGAT 9888
 QY 344 ValPheThrLeuAspAlaAspPhe 351

Search completed: October 30, 2005, 08:56:03
Job time : 185 secs



GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 30, 2005, 08:09:55 ; Search time 661 Seconds
(without alignments)
4385.636 Million cell updates/sec

Title: US-10-698-228-1
Perfect score: 1821
Sequence: 1 MGDPERPEAAGLDQDERSSS.....QRFAVMAETTDNVFTLDADF 351

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9784742 seqs, 4129495052 residues

Total number of hits satisfying chosen parameters: 19569484

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlh
-MODEL=frame+pn2n.model -US10698228/runat 26102005_100716_5225/app_query.fasta_1.519
-Q=/cgn2_1/USPTO_spool/US10698228 -SUFFIX=p2n.rnpb -MINMATCH=0.1
-DB=Published Applications NA -OPMT=fastap -LOOPTXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-LOOPTCL=0 -LOOPTXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40 cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODER=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10698228 @cgn 1 1 354 @runat 26102005_100716_5225
-NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=120 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PTC_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PTC_US_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
19: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:
20: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:
21: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:
22: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:
23: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
24: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
25: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:
26: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
27: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
28: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	100.0	1053	19	US-10-698-228-2
2	1821	100.0	1053	19	US-10-698-228-12
3	1821	100.0	1081	19	US-10-698-228-4
4	1821	100.0	4955	18	US-10-172-118-71
5	1821	100.0	4955	19	US-10-342-887-71
6	1821	100.0	4955	19	US-10-698-228-3
7	1460.5	80.2	1989	9	US-09-925-301-505
8	1460.5	80.2	2216	16	US-10-084-817-342
9	1460.5	80.2	2482	24	US-10-220-335-514
10	1460.5	80.2	2500	9	US-09-954-456-724
11	1460.5	80.2	2500	9	US-09-954-456-1169
12	1460.5	80.2	2500	9	US-09-954-456-1827
13	1460.5	80.2	2500	19	US-10-641-643-1370
14	1460.5	80.2	2500	21	US-10-733-878-458
15	1460.5	80.2	2500	22	US-10-843-641A-3751
16	1460.5	80.2	2500	22	US-10-843-641A-4196
17	1460.5	80.2	2500	22	US-10-843-641A-4854
18	1460.5	80.2	2500	24	US-10-756-149-713
19	1447.5	79.2	2641	24	US-10-220-335-170
20	1443	79.2	2113	22	US-10-764-420-1636
21	1425.5	78.3	1328	19	US-10-403-571-75
22	1376.5	75.6	1371	24	US-10-450-763-15278
23	1348	74.0	1289	26	US-11-097-143-19181
24	1280	70.3	186854	22	US-10-872-156-34
25	1241.5	68.2	3945	26	US-11-097-143-19180
26	1230	67.5	1218	16	US-10-128-714-7245
27	1220.5	67.0	1292	18	US-10-369-493-27809
28	1192.5	65.5	1146	16	US-10-128-714-2245
29	1192	65.5	1242	17	US-10-032-585-6537
30	1176.5	64.6	963	18	US-10-369-493-34052
31	1168.5	64.2	1200	18	US-10-369-493-45857
32	1161.5	63.8	1173	18	US-10-369-493-46252
33	1161.5	63.8	1206	18	US-10-320-797-2145
34	1151	63.2	1314	16	US-10-128-714-1245
35	1151	63.2	1450	16	US-10-128-714-6245
36	1151	63.2	3314	16	US-10-128-714-245
37	1151	63.2	3450	16	US-10-128-714-5245
38	1147.5	63.0	1586	19	US-10-424-599-73024
39	1133.5	62.2	1169	19	US-10-425-114-31361
40	1133.5	62.2	2992	21	US-10-425-115-31182
41	1132.5	62.2	1725	20	US-10-437-963-11109
42	1128.5	62.0	1255	20	US-10-767-701-13723
43	1125.5	61.8	1791	21	US-10-425-115-63321
44	1114.5	61.2	1248	19	US-10-424-599-126843
45	1113	61.1	1146	9	US-09-822-830A-174

ALIGNMENTS

RESULT 1
US-10-698-228-2
; Sequence 2, Application US/10698228
; Publication No. US2004007253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28

```
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-2
Alignment Scores:
Pred. No.: 1
Score: 1821.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Indels: 0
Gaps: 0
DB: 19

US-10-698-228-1 (1-351) x US-10-698-228-2 (1-1053)
QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 1 ATGGGCGACCCGGGAAGCGCGAGCGCGCGGCTGGATCAGATCAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 61 GACACCAACGAAAGTGAATGAATCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysGlnAlaGln 60
DB 121 CGGTTTGTTCATCTTCCAAATCCAGTACCTGATATTTGGAAATGATATAACAGGACAG 180
QY 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
DB 181 GTTCTCTCTGACACGACGAGAGGTGCGACTATCAAGGATCTCCCTACTGGAACAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
DB 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 301 GGAATTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGGTTCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTTCTATGGCTTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
QY 161 ThrMetProTyrValLysLysAlaAspTyrPheAlaLeuArgTyrPheAlaAspArgLys 180
DB 481 ACCATGCCCTATGTTAAGAAAAAAGACAGATTTGGGCTTGGATGGATAGCAGATAGAAA 540
QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluLysValPhePheSerGly 200
DB 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGTGTCTTCTTCAGGA 600
QY 201 SerPheAlaAlaIlePheTyrPheLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 601 TCTTTTGTCTATATCTTGGCTAAGAGAGAGAGGTCTTATGTCAGGACTCATTCTTTC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 661 AATGAACATCATCAGCAGATGAAGGACTTCACTGTGACTTTGCTGCTGTGCTGTGCTCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB 721 TACTTAGTAATAAGCCTTTCAGAAAGAGGCTCAGGAGATCATTTGTGATGCTCTCAAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 781 ATTGAGCAGAGAGTATTTAAACAGAGCCTTCCAGTTGGCCTCATTTGGAATGAATTGCATT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 841 TTGATGAAACAGTACATTTGATGTTGTAGCTGACAGATTTACTTTGGAACTTTGGATCTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 901 AAGGTTTTTCAGGCAGAAAAATCCTTTTGAATTTATGGAACAATTTCTTTTAGAAGAAAA 960
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTGAGTTTATGCGAGAAACC 1020
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1021 ACAGATTAACGCTCTTCACCTTGGATGCAGATTTT 1053

RESULT 2
US-10-698-228-12
; Sequence 12, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-12
Alignment Scores:
Pred. No.: 1
Score: 1821.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Indels: 0
Gaps: 0
DB: 19

US-10-698-228-1 (1-351) x US-10-698-228-12 (1-1053)
QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 1 ATGGGCGACCCGGGAAGCGCGAGCGCGGCTGGATCAGATCAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 61 GACACCAACGAAAGTGAATGAATCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysGlnAlaGln 60
DB 121 CGGTTTGTTCATCTTCCAAATCCAGTACCTGATATTTGGAAATGATATAACAGGACAG 180
QY 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
DB 181 GTTCTCTCTGACACGACGAGAGGTGCGACTATCAAGGATCTCCCTACTGGAACAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
DB 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 301 GGAATTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGGTTCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTTCTATGGCTTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
QY 161 ThrMetProTyrValLysLysAlaAspTyrPheAlaLeuArgTyrPheAlaAspArgLys 180
DB 481 ACCATGCCCTATGTTAAGAAAAAAGACAGATTTGGGCTTGGATGGATAGCAGATAGAAA 540
QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluLysValPhePheSerGly 200
DB 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGTGTCTTCTTCAGGA 600
QY 201 SerPheAlaAlaIlePheTyrPheLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 601 TCTTTTGTCTATATCTTGGCTAAGAGAGAGAGGTCTTATGTCAGGACTCATTCTTTC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 661 AATGAACATCATCAGCAGATGAAGGACTTCACTGTGACTTTGCTGCTGTGCTGTGCTCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB 721 TACTTAGTAATAAGCCTTTCAGAAAGAGGCTCAGGAGATCATTTGTGATGCTCTCAAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 781 ATTGAGCAGAGAGTATTTAAACAGAGCCTTCCAGTTGGCCTCATTTGGAATGAATTGCATT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 841 TTGATGAAACAGTACATTTGATGTTGTAGCTGACAGATTTACTTTGGAACTTTGGATCTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 901 AAGGTTTTTCAGGCAGAAAAATCCTTTTGAATTTATGGAACAATTTCTTTTAGAAGAAAA 960
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTGAGTTTATGCGAGAAACC 1020
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1021 ACAGATTAACGCTCTTCACCTTGGATGCAGATTTT 1053
```

Db 181 GCTTCCTCTTGGACACAGAGGTTGACTTATCAAGAGTCTCCCTCAGTCAAGCAAG 240
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAaSerAsp 100
Db 241 CTTAAGCAGATGAGAGTACTTCTCATCTCTCATCTTACCTTTTGGAGCAGTGAT 300
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 301 CGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTCCAGAGGCT. 360
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 361 CGCTGTTTCTATGGCTTTCAATTTCTCATGAGAATTTTCTCAGAGATGTACAGTTG 420
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 421 CTGATAGACATTTACATCAGATCCCAAGAAAGGAATTTTATTTAATGCAATGAA 480
Qy 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGCCCTTGGATGGATAGCAGATAGAAA 540
Qy 181 SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 200
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGGCTGTGTAGAGAGTCTTCTCTCAGGA 600
Qy 201 SerPheAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGGCTGTATATCTGGCTTAAAGAAAGAGAGGCTTATGTCAGGAGCTCACTTTTCC 660
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 ATGAACTCATCAGCAGATGAGAGCTTCTCTGCTGTAGAGAGTCTTCTCTCAGGA 720
Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 721 TACTTAGTAAATAAGCTTTCAGAAAGAGGTCAGGAGATCATTTGTGTGCTGTCAAA 780
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 781 ATTGAGCAGAGTCTTAAACAGAGGCTTGGCCAGTGGCCCTCATTTGGAATGCAAT 840
Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 841 TTGATGAACAGTACATTCAGTTGTAGCTGTAGCAGATTCCTTGTGAACTTGGATTTCTCA 900
Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 901 AAGGTTTTTCAGCAGAGAAATCCTTTTGAATTTATGGAACACATTTCTTTAGAGGAAA 960
Qy 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 961 ACAAAATTTCTTTCAGAAACAGTTCAGAGTATCAGCGTTTTCAGTATTCAGTATTCAGAAACC 1020
Qy 341 ThrAsnAsnValPheThrLeuAspAlaAspPhe 351
Db 1021 ACAGATAACGTCTTCACCTTGGATGAGATTTT 1053

RESULT 3

US-10-698-228-4
; Sequence 4, Application US/10698228
; Publication No. US2004007253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391

; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4

; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-4

Alignment Scores:
Pred. No.: 1,62e-216 Length: 1081
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 19 Gaps: 0

US-10-698-228-1 (1-351) x US-10-698-228-4 (1-1081)

Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 20 ATGGCGCACCCGGAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 79
Qy 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
Db 80 GACACCAACGAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGATTTCTCGC 139
Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTTrpLysMetTyrLysGlnAlaGln 60
Db 140 CGGTTTGTGATCTTTCCAAATCCAGTACCTCATATTTGGAATAATGATAAAGCAGGACAG 199
Qy 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db 200 GCTTCTCTTGGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCAGTGGAAACAG 259
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 260 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGCCAGCTGAT 319
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 320 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 379
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 380 CGCTGTTTCTATGGCTTTCAAATTTCTATCAGAGAAATGTTCACTCAGAGATGTACAGTTTG 439
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 499
Qy 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 500 ACCATGCCCTTATGTTAAGAAAAAAGCAGATTGGCCCTTGGATGGATAGCAGATAGAAA 559
Qy 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db 560 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGAGTCTTCTCTCAGGA 619
Qy 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 620 TCTTTTGGCTGTATATCTTGGCTTAAAGAAAGAGAGGCTCTTATGCCAGGACTCACTTTTCC 679
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 680 AATGAACTCATCAGCAGAGATGAAGGACTTTCACCTGTGACTTTGTGCTGCTGATGTTCAA 739
Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 740 TACTTAGTAAATAAGCCTTTCAGAAAGAGGCTCAGGAGATCATTTGTTGCTGTCTCAAA 799

```
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 800 ATGTAGCAGAGGTTTAAACAGAGCCTTGCCAGTTGGCCTCATTTGAATGAATTCATT 859
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 860 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 919
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 920 AAGGTTTTTCAGCAGAAAATCCTTTGATTTTATGGAATAACATTTCTTTAGAAAGGAAA 979
QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 980 ACAAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1039
QY 341 ThrAsnValPheThrLeuAspAlaAspPhe 351
Db 1040 ACAGATAACGTCTTCACCTTGGATGCAGATTTT 1072

RESULT 4
US-10-172-118-71
; Sequence 71, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AB036063
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-71

Alignment Scores:
Pred. No.: 1,77e-215 Length: 4955
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-10-698-228-1 (1-351) x US-10-172-118-71 (1-4955)
QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 245 ATGGGCGACCCGGAAGGCGGAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 304
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
Db 305 GACACCAACGAAGTGAATTAAGTCAATGATGAGACGCTCTCTAAGAAAGAGTTCTCGC 364
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Db 365 CGGTTTGTCTATCTTCCATCCAGTACCTGATATTGGAAAATGTATAAACAGGCACAG 424
QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
```

```
Db 425 GCTTCCTCTGGACAGCAGAGGTCGACTTATCAAAAGGATCTCCCTCACTGGAACAAG 484
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAspSer 100
Db 485 CTTAAAGCAGATGAGAAGTACTTTCATCTCTACATCTTAGCCTTTTTCAGCAGCAGTAT 544
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGTTTCCAGAGCT 604
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 605 CGCTGTTTCTATGGCTTTCAAATTCATCAGAGATGTTTCTCAGAGATGTACAGTTTG 664
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 724
QY 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 725 ACCATGCCCTTATGTTAAGAAAAGCAGATTGGGCCCTTGGATGGATAGCAGATAGAAA 784
QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGluGlyValPhePheSerGly 200
Db 785 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 844
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 845 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACATTTTTC 904
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 905 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTCCA 964
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 965 TACTTAGTAATAAGCCTTCAGAAAGAGGTTCAGGAGATCATTTGTGCTGTCTCA 1024
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 1025 ATTGAGCAGGAGTTTTTAAACAGAAAGCCTTCCAGTTGGCCTCATTTGCAATTCATT 1084
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 1085 TTGATGAACAGTACATTTGATGTTGTAGCTGACAGATTACTTTGGAACTTGGATTTCCA 1144
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 1145 AAGGTTTTTCAGCAGAAAATTCCTTTGATTTTATGGAATAACATTTCTTTAAGAGGAAA 1204
QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 1205 ACAAAATTTCTTTCAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1264
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1265 ACAGATAACGTCTTTCACCTTGGATGCAGATTTT 1297

RESULT 5
US-10-342-887-71
; Sequence 71, Application US/10342887
; Publication NO. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
```

Qy	241	Tyrl	ValAsnLysProSerGluArgValargGluileleValAspAlaVallys	260
Db	965	TACT	TAGTAAATAAGCCCTTCAGAAGAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA	1024
Qy	261	IleGlu	GlnGluPheIeThrGluAlaLeuProValGlyLeuileGlyMetAsnCysIle	280
Db	1025	ATTG	AGCAGAGGATTTTTAAACAGAAAGCCCTGCCAGTTGGCTCTCAATTGGAATGAATTCGATT	1084
Qy	281	LeuMet	LysGlnTyrlleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer	300
Db	1085	TTGAT	GAACACAGTACATTGAGTTTGTAGCTGCAGAGATTACTTGTGGNACTTGGATTCTCA	1144
Qy	301	LysVal	PheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlylys	320
Db	1145	AAGG	TTTTTTCAGGCAGAGAAATCCTTTTGAATTTATATGAAAAACATTTCTTTAGAAAGGAAAA	1204
Qy	321	ThrAsn	PhePheGluLysArgValSerGluTyrlGlnArgPheAlaValMetAlaGluThr	340
Db	1205	ACAA	ATTTCTTTCAGAAAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC	1264
Qy	341	ThrAsp	AsnValPheThrLeuAspAlaAspPhe	351
Db	1265	ACAG	TAACGCTTTCACCTTGGATGCAGATTTT	1297
RESULT 6				
US-10-698-228-3				
; Sequence 3, Application US/10698228				
; Publication No. US20040072253A1				
; GENERAL INFORMATION:				
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA				
; TITLE OF INVENTION: New protein and its DNA				
; FILE REFERENCE: 2619WOOP				
; CURRENT APPLICATION NUMBER: US/10/698,228				
; CURRENT FILING DATE: 2003-10-30				
; PRIOR APPLICATION NUMBER: US/10/019,733				
; PRIOR FILING DATE: 2001-12-28				
; PRIOR APPLICATION NUMBER: JP 11-181131				
; PRIOR FILING DATE: 1999-06-28				
; PRIOR APPLICATION NUMBER: JP 11-192391				
; PRIOR FILING DATE: 1999-07-06				
; PRIOR APPLICATION NUMBER: JP 2000-017770				
; PRIOR FILING DATE: 2000-01-21				
; NUMBER OF SEQ ID NOS: 14				
; SEQ ID NO 3				
; LENGTH: 4955				
; TYPE: DNA				
; ORGANISM: Homo sapiens				
; FEATURE:				
; OTHER INFORMATION:				
US-10-698-228-3				
Alignment Scores:				
Pred. No.: 1,77e-215 Length: 4955				
Score: 1821.00 Matches: 351				
Percent Similarity: 100.00% Conservative: 0				
Best Local Similarity: 100.00% Mismatches: 0				
Query Match: 100.00% Indels: 0				
DB: 19 Gaps: 0				
US-10-698-228-1 (1-351) x US-10-698-228-3 (1-4955)				
Qy	1	MetGly	AspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer	20
Db	245	ATGG	CGACCCCGAAAGCGCGAAGCGCGCGCTGGATCAGATGAGATCATCTTCA	304
Qy	21	AspThr	AsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg	40
Db	305	GACAC	CAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTTCG	364
Qy	41	ArgPhe	ValIlePheProIleGlnTyrlProAspIleTyrlPysMetTyrlLysGlnAlaGln	60
Db	365	CGGT	TTGTGTCATCTTCCAAATCCAGTACCTGTATTTGGAAAAATGTATAAACAGGCAC	424

```
QY 61 AlaSerPheThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
DB 425 GCTTCTCTCTGACACAGAGAGGTGAGTCTATCAAGAGTCTCTCTCTGAGCAAG 484
QY 81 LeuLysAlaAspGluLysTrpPheLysHisLeuLeuAlaPheAlaAsp 100
DB 485 CTTAAAGCAGATGAGAAGTACTTCTCTCTACATCTTAGCCCTTTTTCAGCCAGTAT 544
QY 101 GlyLeuValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 545 CGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGCT 604
QY 121 ArgCysPheThrGlyPheGlnLeuLeuLeuGluAsnValHisSerGluMetTrpSerLeu 140
DB 605 CGCTGTTTCTATGCTCTTCAATTCATTCGAGATGTTCTCTCAGAGATGATACAGTTTG 664
QY 141 LeuLeuAspThrTrpLeuArgAspProLysLysArgGluPheLeuPheAsnAlaLeuGlu 160
DB 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 724
QY 161 ThrMetProTrpValLysLysLysAlaAspTrpAlaLeuArgTrpLeuAlaAspArgLys 180
DB 725 ACCATGCCCTATGTTAAAGAAAAAGCAGATTTGGCCCTTGGCATGGATAGCAGATAGAAA 784
QY 181 SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 200
DB 785 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTAGAGAGGATTTTCTCTCAGGA 844
QY 201 SerPheAlaAlaLeuPheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 845 TCTTTTGTCTATATTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 904
QY 221 AsnGluLeuLeuSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 905 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluLeuLeuValAspAlaValLys 260
DB 965 TACTTAGTAAATAAGCTTTTCAAGAAAGGGTCAAGGAGATCATTTGTTGCTGTCAAA 1024
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuLeuGlyMetAsnCysIle 280
DB 1025 ATTGACGAGAGTTTTACAGAGAGCTTGGCAGTTGGCTCATTTGAGATGAATTCAT 1084
QY 281 LeuMetLysGlnTrpIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 1085 TTGATGAACAGTACATTCAGTTGTAGCTGACAGATTTACTTGTGAACTTGGATTCTCA 1144
QY 301 LysValPheGluAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 1145 AAGTTTTTTCAGCAGAAAATCCTTTGATTTTATGGAAGAAATTTCTTTAGAGGAAAA 1204
QY 321 ThrAsnPhePheGluLysArgValSerGluTrpGlnArgPheAlaValMetAlaGluThr 340
DB 1205 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTTATGGCAGAAACC 1264
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1265 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1297
```

RESULT 7

```
US-09-925-301-505
; Sequence 505, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT FILING DATE: 2001-08-10
; PRIOR FILING DATE: 2000-03-08
```

```
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 505
; LENGTH: 1989
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1917)
; OTHER INFORMATION: n equals a.t.g, or c
US-09-925-301-505
Alignment Scores:
Pred. No.: 4.56e-171 Length: 1989
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 9 Gaps: 1
```

US-10-698-228-1 (1-351) x US-09-925-301-505 (1-1989)

```
QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
DB 185 CTTGAGCGGAGACCGCGCTCTCTGGCCAGCAAGACCGCGAGGAGGATCTTCCAGGAGCCAC 244
QY 23 nGluSerGluLeuLysSer-----AsnGluGluProLeuLeuArgLysE 38
DB 245 GGAGCGCAAACTAAAGCAGCTGCCCGCGGTGGAGGATGAGCGCTGCTCAGAGAAAA 304
QY 38 rSerArgArgPheValIlePheProIleGlnTrpProAspIleTrpLysMetTrpLysG 58
DB 305 CCCCCCGCTTTGTCTATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 364
QY 58 nAlaGlnAlaSerPheTrpAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
DB 365 GGACAGAGCTTCTTTTGGACCGCGGAGGAGGTGGACCTTCCAGAGGACATTCAGCATG 424
QY 78 pAsnLysLeuLysAlaAspGluLysTrpPheLysHisLeuAlaPheAlaAl 98
DB 425 GGAATCCCTGAAACCCCGAGGAGAGATATTTATATCCATGTTCTGGCTTCTTTGCGAGC 484
QY 98 aSerAspGlyLeuValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
DB 485 AAGCGATGGCATAGTAATAAGAACTTTGGTGGAGCGATTTAGCCAAAGAAATTCAGATTAC 544
QY 118 oGluAlaArgCysPheTrpGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
DB 545 AGAAGCCCGCTGTTCTATGGCTTCCAAATTTGCCATGGAAAAACATACATTTCTGAANTGT 604
QY 138 rSerLeuLeuLeuAspThrTrpIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
DB 605 TAGTCTCTTATTGACACTTACATAAAAGATCCCAAGAAAGGAAATTTCTTCTCAATGC 664
QY 158 alLeGluThrMetProTrpValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
DB 665 CATTGAAACGATGCTTGTGTCAAGAAAGAGGAGACTGGGCGCTTGGCGCTGGATTTGGGGA 724
QY 178 pArgLysSerThrPheGlyGluArgValAlaAlaPheAlaAlaValGluGlyValPhePh 198
DB 725 CAAAGAGGCTACCTATGTTGAAACGTGTTGTAGCCCTTTGCTCAGTGGAGGCAATTTCTT 784
QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
DB 785 TTCGGTCTTTTGGTTCGATATTTCTGGCTCAAGAAACGAGACTGATGCTGGCGCTCAC 844
QY 218 rPheSerAsnGluLeuLeuSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
DB 845 ATTTTCTAATGAACCTTATTAGCAGAGATGAGGGTTTACACTGTGATTTTGTGCTGAT 904
QY 238 tPheGlnTrpLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
```


Db 302 CCTGAGCGGACCCCGCTCTGCGCCAGCAAGACCGCGAGGAGATCTTCCAGGAGCCAC 361
Qy 23 nGluSerGluLeuLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 362 GGAGCGGAAACTAAAGCAGCTCCCGCGCGTGGAGGATGAGCGCTGCTGAGAGAAA 421
Qy 38 rSerArgArgPheValIlePhePheProIleGlnTyrProAspIleTrrPlysMetTrrLysGl 58
Db 422 CCCCCCGCGCTTGTCTATCTCCCATCAGTACCATGATATCTGGCAGATGTATAAGAA 481
Qy 58 nAlaGlnAlaSerPheTrrPheAlaGluValAspLeuSerLysAspLeuProHisTr 78
Db 482 GGAGAGGCTTCTTTTGACCGCCGAGGAGGTGACCTTCCAGGACATTCAGCACTG 541
Qy 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98
Db 542 GGAATCCCTGAAACCCGAGGAGATATTTATATCCCATGTTCTGGCTTCTTGGAGC 601
Qy 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 602 AAGCGATGCGATAGTAATAAGAACTTGTGTGAGCGATTTAGCCAGAGTTCAGATTAC 661
Qy 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 662 AGAAGCGCGCTGTTCTATGGCTTCCAAATTCCTGCAAAACATACATCTCTGAATGTA 721
Qy 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db 722 TAGTCTCTTATGACACTTACATAAAGATCCCAAGAAAGGGAATTTCTTCTCAATGC 781
Qy 158 aIleGluThrMetProTyrValLysLysLysAlaAspTrrPheAlaLeuArgTrrPheAlaAs 178
Db 782 CATTGAACGATGCTTGTGTCAAGNAGAGGAGAGCTGGCGCTTGGCTGGAGTGGGA 841
Qy 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
Db 842 CAAAGAGGCTACTTGTGTGAACGTGTGTAGCCTTGTGCTGAGTGGAGGCAATTTCTT 901
Qy 198 eSerGlySerPheAlaAlaIlePheTrrPheLysLysArgGlyLeuMetProGlyLeuTh 218
Db 902 TTCCGGTCTTGTGGTTCGATATTTCTGGCTCAAGAAACGAGGACTGATGCTGGCCTCAC 961
Qy 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db 962 ATTTCTAATGACTTATTAGCAGATGAGGTTTACACTGTGATTTTCTTGGCTGTAT 1021
Qy 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db 1022 GTTCAACACCTGTTACACAAACCATCGGAGGAGAGTAAGAGAAATATTAATCAATGC 1081
Qy 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db 1082 TGTTCCGATAGAACAGGAGTTCTCTCACTGAGCGCTTGCCTGTGAAGCTCATTTGGATGAA 1141
Qy 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGl 298
Db 1142 TTGCACTCTAATGAAGCAATACATTGAGTTTGTGGCAGACACACTTATCTGGAACCTGG 1201
Qy 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
Db 1202 TTTTAGCAAGGTTTTCAGAGTAGAGNACCCATTTGACTTTATGGAGAAATATTCTACTGA 1261
Qy 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
Db 1262 AGGAAAGACTAACTTCTTTGAGAAGAGAGTAGGCGAGTACAGAGATGGGAGTGTATGTC 1321
Qy 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1322 AAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 11

US-09-954-456-1169

; Sequence 1169, Application US/0995456

; Patent NO. US20020115057A1

; GENERAL INFORMATION:

; APPLICANT: Young, Paul

; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can

; TITLE OF INVENTION: Sets

; FILE REFERENCE: 689290-76

; CURRENT FILING DATE: 2001-09-18

; PRIOR APPLICATION NUMBER: US/60/233,617

; PRIOR FILING DATE: 2000-09-18

; PRIOR APPLICATION NUMBER: US/60/234,052

; PRIOR FILING DATE: 2000-09-20

; PRIOR APPLICATION NUMBER: US/60/234,923

; PRIOR FILING DATE: 2000-09-25

; PRIOR APPLICATION NUMBER: US/60/235,134

; PRIOR FILING DATE: 2000-09-25

; PRIOR APPLICATION NUMBER: US/60/235,637

; PRIOR FILING DATE: 2000-09-26

; PRIOR APPLICATION NUMBER: US/60/235,638

; PRIOR FILING DATE: 2000-09-26

; PRIOR APPLICATION NUMBER: US/60/235,711

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: US/60/235,720

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: US/60/235,840

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: US/60/235,863

; PRIOR FILING DATE: 2000-09-27

; NUMBER OF SEQ ID NOS: 2276

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 1169

; LENGTH: 2500

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-09-954-456-1169

Alignment Scores:
Pred. No.: 6,53e-171 Length: 2500
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 9 Gaps: 1

US-10-698-228-1 (1-351) x US-09-954-456-1169 (1-2500)

Qy 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
Db 302 CCTGAGCGGACCCCGCTCTGCGCCAGCAAGACCGCGAGGAGATCTTCCAGGAGCCAC 361
Qy 23 nGluSerGluLeuLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 362 GGAGCGGAAACTAAAGCAGCTGCGCGCGTGGAGGATGAGCGCTGCTGAGAGAAA 421
Qy 38 rSerArgArgPheValIlePhePheProIleGlnTyrProAspIleTrrPlysMetTrrLysGl 58
Db 422 CCCCCCGCGCTTGTCTATCTCCCATCAGTACCATGATATCTGGCAGATGTATAAGAA 481
Qy 58 nAlaGlnAlaSerPheTrrPheAlaGluValAspLeuSerLysAspLeuProHisTr 78
Db 482 GGAGAGGCTTCTTTTGACCGCCGAGGAGGTGACCTTCCAGGACATTCAGCACTG 541
Qy 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98
Db 542 GGAATCCCTGAAACCCGAGGAGATATTTATATCCCATGTTCTGGCTTCTTGGAGC 601
Qy 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 602 AAGCGATGCGATAGTAATAAGAACTTGGTGGAGCGGATTTAGCCAGAGATTCAGATTAC 661
Qy 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 662 AGAAGCGCGCTGTTCTATGGCTTCCAAATTCCTGCAAAACATACATCTCTGAATGTA 721

QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1322 AAGTCAACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 14

US-10-733-878-458
; Sequence 458, Application US/10733878
; Publication No. US20040224408A1
; GENERAL INFORMATION:
; APPLICANT: Jean-Philippe Girard
; APPLICANT: Francois Amalric
; APPLICANT: Myriam Roussigne
; APPLICANT: Thomas Clouaire
; TITLE OF INVENTION: THAP PROTEINS AS NUCLEAR RECEPTORS FOR
; TITLE OF INVENTION: CHEMOKINES AND ROLES IN TRANSCRIPTIONAL REGULATION, CELL
; TITLE OF INVENTION: PROLIFERATION AND CELL DIFFERENTIATION
; FILE REFERENCE: BIOBANK.012A
; CURRENT APPLICATION NUMBER: US/10/733,878
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 60/432699
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/485027
; PRIOR FILING DATE: 2003-07-03
; NUMBER OF SEQ ID NOS: 535
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 458
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-733-878-458

Alignment Scores:

Pred. No.: 6,53e-171 Length: 2500
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 21 Gaps: 1

US-10-698-228-1 (1-351) x US-10-733-878-458 (1-2500)

QY 4 ProGluArgProGluAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
Db 302 CTGAGCGGAGCCCGCTCTGCCAGCAGACCGGAGAGGATCTTCAGAGGCCAC 361
QY 23 nGluSerGluLeuLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 362 GGAGCGGAAACTAAAGCAGCTGCCCGCGGTGGAGGATGAGCCGCTGCTGAGAGAAA 421
QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGI 58
Db 422 CCCCCCGCGCTTGTCTATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481
QY 58 nAlaGlnAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisStr 78
Db 482 GGCAGAGGCTTCTTTTGGACCGCCGAGGAGGTGTACCTTCCAGACATCATCAGCACTG 541
QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98
Db 542 GGNATCCCTGAAACCCGAGAGAGATATTTATATCCATGTTCTGGCTTCTTTCGAGC 601
QY 98 aSerAspGlyLeValAsnGlnAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 602 AAGCGATGCGATAGTAATAATGAAACTTGTGGAGCGATTAGCCAAAGAGTTCCAGATTAC 661
QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 662 AGAAGCCCGCTGTTCTATGGCTTCCAAATTTGCCATGGAAACATACATCTCGAAATGTA 721
QY 138 rSerLeuLeuLeuAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db 722 TAGTCTTCTTATTCACACTTACATATAAAGNATCCCAAGAAAGAGGAATTTCTTCAATGC 791

QY 158 aileGluThrMetProTyrValLysLysAlaAspTIPAlaLeuArgTIPilleAlaAs 178
Db 782 CAITGAACGATGCCCTTGTCTCAAGAAAGAGGAGACTGGGCTTGGCTGGATGGGA 841
QY 178 pArgLysSerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePh 198
Db 842 CAAAGAGGCTACCTATGTTGAACGCTGTGTAGCCTTTGCTGACGTGGAAGGCATTTCTT 901
QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db 902 TTCCGGTTCCTTTTGGCTGCATATTTGCTCAAGAAACGAGGACTGATGCTGGCCTCAC 961
QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db 962 ATTTTCTAATGAACCTTATTAGCAGAGATGAGGTTTACACTGTGATTTTGTGCTGCTGAT 1021
QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db 1022 GTTCAAAACACCTGGTACACAAACCATCGAGGAGAGTAGAGAGAAATAATTATCAATGC 1081
QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db 1082 TGTTCGGATAGAACAGGAGTTCTCTACTGAGGCCCTTGCTGTGAAGCTCATTTGGGATGAA 1141
QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGl 298
Db 1142 TTGCACCTTAATGAACATACATTTGAGTTTGTGGCAGACAGACTTATGCTGGAACTGGG 1201
QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
Db 1202 TTTTAGCAAGGTTTTTCAGAGTAGAGAACCCATTTGACTTTTATGGAGAAATATTTCACTGGA 1261
QY 318 uGlyLysThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
Db 1262 AGGAAAGACTAACTTCTTTGAGAAGAGAGTAGGCGAGTATCAGAGGATGGGAGTGATGTC 1321
QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1322 AAGTCAACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 15

US-10-843-641A-3751
; Sequence 3751, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; TITLE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3751
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-3751

Alignment Scores:
Pred. No.: 6,538-171 Length: 2500
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 22 Gaps: 1

US-10-698-228-1 (1-351) x US-10-843-641A-3751 (1-2500)

```
QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
DB 302 CCTGAGCGGGACCCGCGCTCTGGCAGCAGACCGCGGAGGAGATCTCCAGGAGCCAC 361
QY 23 nGluSerGluLeuLysSer-----AsnGluGluProLeuLeuArgLysSe 38
DB 362 GGAGCCGAAACTAAAGCAGCTGCCCGCGCGTGGAGGATGAGCGCTGCTGAGAGAAAA 421
QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGl 58
DB 422 CCCCGCGCGCTTGTTCATCTCCCATCGAGTACCATGATATCTGCGCAGATGATAAGAA 481
QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
DB 482 GCGAGAGGCTTCCTTTTGACCGCGCGAGGAGTTGACCTCTCCAAGGACATTCAGCAGCTG 541
QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98
DB 542 GGAATCCCTGAAACCGAGGAGAGATATTTATATCCCATGTTCTGGCTTCTTTCGACG 601
QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
DB 602 AAGCGATGGCATAGTAATGAAACTTGGTGGAGCGATTTAGCCAAAGAAAGTTTCAGATTAC 661
QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
DB 662 AGAAGCCCGCTGTTCATGGCTTCCAAATGCCATGGAAACATACATCTCGAAATGTA 721
QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
DB 722 TAGTCTTCTTATTGACACTTACATNAAGATCCCAAGAAAGGGAATTTCTTTCATGTC 781
QY 158 aIleGluThrMetProTyrValLysLysAlaAspTrpAlaLeuA:GTrpIleAlaAs 178
DB 782 CATTAAGACGATGCTTGTGTCAAGAAAGACGACCTGGGCGCTTGGCTGGATTGGGA 841
QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePh 198
DB 842 CAAGAGGCTACCTATGTGTGAACGTTGTGTAGCCCTTGTGTCAGTGGAGGCAATTTCTT 901
QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
DB 902 TTCGGTTCCTTTGGCTCGATATCTGGCTCAAGAAACGAGGACATGATGCGCTGCCTCAC 961
QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
DB 962 ATTTTCTAATGAATTAATAGCAGAGATGAGGGTTTACATGTGATTTTGTCTGCTGAT 1021
QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
DB 1022 GTTCAACACCTGGTGTACCAACACCATCGAGGAGAGTAAAGAGAAATAATTTATCAATGC 1081
QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
DB 1082 TGTTCGGATAGACAGGAGTTCCTCCTGAGGCCCTTGTGTAAGCTCATTGGGATGAA 1141
```

```
QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG1 298
DB 1142 TTGCACCTCTAATGAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACTGGG 1201
QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
DB 1202 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCATTGACTTTATGAGAGAAATATTTCACTGGA 1261
QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
DB 1262 AGGAAAGACTTAACCTCTTTGAGAAAGAGAGTAGGCGAGTATCAGAGGATGGGAGTGATGTC 1321
QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1322 AAGTCCACACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361
```

Search completed: October 30, 2005, 10:18:11
Job time : 679 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2005, 21:42:15 ; Search time 3258 Seconds
(without alignments)
15660.940 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1053

Sequence: 1 atggcgaccgcgaaagcc.....tcaccttgatgcagatattt 1053

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1053	100.0	1053	6	AR454868 Sequence
2	1053	100.0	1053	6	BD064764 Novel pro
3	1053	100.0	1053	6	BD093077 Novel pro
4	1053	100.0	1056	6	CQ714252 Sequence
5	1053	100.0	1056	9	AB166671 Homo sapi
6	1053	100.0	4955	6	AR454869 Sequence
7	1053	100.0	4955	6	BD064765 Novel pro
8	1053	100.0	4955	6	BD093078 Novel pro
9	1053	100.0	4955	9	AB036063 Homo sapi
10	1051.4	99.8	1053	6	AR454877 Sequence
11	1051.4	99.8	1053	6	BD064773 Novel pro
12	1051.4	99.8	1053	6	BD093086 Novel pro
13	1051.4	99.8	1081	6	AR454870 Sequence
14	1051.4	99.8	1081	6	BD064766 Novel pro
15	1051.4	99.8	1081	6	BD093079 Novel pro
16	1049.8	99.7	1601	6	BD156916 Primer fo
17	1049.8	99.7	1601	6	AX877905 Sequence
18	1049.8	99.7	1601	9	AK001965 Homo sapi
19	895.2	85.0	4532	10	BC058103 Mus muscu

20	855	81.2	4571	9	BC042468	BC042468 Homo sapi
21	853.4	81.0	900	9	AB163437	AB163437 Homo sapi
22	837.6	79.5	171737	10	AC122379	AC122379 Mus muscu
23	579	55.0	2500	6	AR300103	AR300103 Sequence
24	579	55.0	2500	6	AR438503	AR438503 Sequence
25	579	55.0	2500	6	AR562011	AR562011 Sequence
26	578.4	54.9	1649	9	BC001886	BC001886 Homo sapi
27	578.4	54.9	1653	9	BC030154	BC030154 Homo sapi
28	578.4	54.9	1794	9	S40301	S40301 Homo sapien
29	578.4	54.9	2500	6	AR380825	AR380825 Sequence
30	578.4	54.9	2500	6	AX333242	AX333242 Sequence
31	578.4	54.9	2500	6	AX333687	AX333687 Sequence
32	578.4	54.9	2500	6	AX334345	AX334345 Sequence
33	578.4	54.9	2500	9	HSRR25S	X59618 H.sapiens R
34	578.4	54.9	3039	9	AK092671	AK092671 Homo sapi
35	578.2	54.9	1562	5	BC061353	BC061353 Xenopus t
36	574.6	54.6	1538	5	CR762018	CR762018 Xenopus t
37	574.6	54.6	1611	5	CR848103	CR848103 Xenopus t
38	574.6	54.6	1647	5	BC080161	BC080161 Xenopus t
39	573.6	54.5	217774	2	AC120603	AC120603 Rattus no
40	573.2	54.4	3203	9	AK123010	AK123010 Homo sapi
41	570.4	54.2	2111	10	MUSRNM2A	M14223 Mouse ribon
42	568.8	54.0	1170	6	E14825	E14825 cDNA encodi
43	566.2	53.8	1570	5	BC072071	BC072071 Xenopus l
44	562.6	53.4	2057	5	BC044355	BC044355 Danio rer
45	561.8	53.4	1406	5	CR353642	CR353642 Gallus ga

ALIGNMENTS

RESULT 1	AR454868	AR454868	Sequence 2 from patent US 6682917.	1053 bp	DNA	linear	PAT 20-FEB-2004
LOCUS	AR454868	Sequence 2 from patent US 6682917.					
DEFINITION	AR454868	Sequence 2 from patent US 6682917.					
ACCESSION	AR454868	Sequence 2 from patent US 6682917.					
VERSION	AR454868.1	GI:42688823					
KEYWORDS	Unknown.						
SOURCE	Unknown.						
ORGANISM	Unknown.						
REFERENCE	1 (bases 1 to 1053)						
AUTHORS	Nakamura, Y., Arakawa, H. and Tanaka, H.						
TITLE	Protein having a ribonucleotide Reductase activity and a DNA thereof						
JOURNAL	Patent: US 6682917-A 2 27-JAN-2004;						
FEATURES	Location/Qualifiers						
source	1..1053						
ORIGIN	/organism="unknown"						
	/mol_type="genomic DNA"						

Query Match	100.0%;	Score 1053;	DB 6;	Length 1053;
Best Local Similarity	100.0%;	Pred. No. 1.2e-255;		
Matches 1053;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	ATGGCGACCCGAAAGCGCGGCTGGATCAGGATGAGATGAGATCATCTTCA	60	
Db	1	ATGGCGACCCGAAAGCGCGGCTGGATCAGGATGAGATGAGATCATCTTCA	60	
Qy	61	GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCCTAAGAAAGATTCTCGC	120	
Db	61	GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCCTAAGAAAGATTCTCGC	120	
Qy	121	CGTTTGTTCATCTTCCCAATCCAGTACCTCGATATTTGGAAAATGTATAAAGCAGCAG	180	
Db	121	CGTTTGTTCATCTTCCCAATCCAGTACCTCGATATTTGGAAAATGTATAAAGCAGCAG	180	
Qy	181	GCTTCCTTTGGACAGCAGAGAGGTCGATATCAAAAGGATCTCCCTCACTGGAAACAAG	240	
Db	181	GCTTCCTTTGGACAGCAGAGAGGTCGATATCAAAAGGATCTCCCTCACTGGAAACAAG	240	
Qy	241	CTTAAAGCAGATGAGAGTACTTCTCATCTTAGCCTTTTGGAGCCAGCATGAT	300	

[illegible]

RESULT	2
LOCUS	BD064764
DEFINITION	Novel protein and DNA thereof.
ACCESSION	BD064764
VERSION	BD064764.1 GI:22610367
KEYWORDS	JP 2001269184-A/1.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 1053)
AUTHORS	Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE	Novel protein and DNA thereof
JOURNAL	Patent: JP 2001269184-A 1 02-OCT-2001;
	YUSUKE NAKAMURA,TAKEDA CHEMICAL INDUSTRIES LTD
COMMENT	OS Homo sapiens (human)

PN	JP 2001269184-A/1
PD	02-OCT-2001
PF	27-JUN-2000 JP 2000192401
PI	YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
	C12N15/09, A61K31/711, A61K38/00, A61K38/43, A61K45/00, A61K48/00, PC
	A61P35/00,
PC	A61P43/00; C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC
	, C12N5/10, C12N9/02,
PC	C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53// (C12N15/09,
PC	C12R1:91),
PC	(C12N15/09, C12R1:19), (C12N9/02, C12R1:91), (C12P21/02, C12R1:91),
PC	(C12P21/08, C12R1:91), C12N15/00, A61K37/02, A61K37/48, C12N5/00,
PC	(C12N15/00, C12R1:91), (C12N15/00, C12R1:19)
CC	Novel protein and DNA thereof
FH	Key
FT	Location/Qualifiers
FT	source 1..1053
FT	Location/Qualifiers /organism='Homo sapiens (human)'.
FEATURES	
source	1..1053
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
ORIGIN	
Query Match	100.0%; Score 1053; DB 6; Length 1053;
Best Local Similarity	100.0%; Pred. No. 1.2e-255;
Matches 1053;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 ATGGGGCACC CGGAAGCGCGGAGCGCGCGCTGGATCAGATCAGAGATCATCTTCA 60
Db	
Qy	1 ATGGGGCACC CGGAAGCGCGGAGCGCGCGCTGGATCAGATCAGAGATCATCTTCA 60
Db	
Qy	61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Db	
Qy	121 CGGTTTGTCA TCTTTCCAAATCCAGTACCGCTGATATTTGGAAATGATATAACAGGCACAG 180
Db	
Qy	181 GCTTCCCTCTGGACAGCAGAGAGGTGCACTTATCAAGAGATCTCCCTCACTGGAAACAAG 240
Db	
Qy	241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTTAGCCCTTTTGTGACGCAGGTAT 300
Db	
Qy	301 GGAATTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Db	
Qy	361 CGCTGTTTCTATGGCTTTCAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db	
Qy	421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTTAATGCAATTGAA 480
Db	
Qy	481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCCTTGGATGGATAGCAGATAGAAAA 540
Db	
Qy	541 TCTACTTTTGGGGAAAGAGTGTGGCTTTGCTGCTGTGAAGAGGAGTTTCTTCTCAGGA 600
Db	
Qy	601 TCTTTTGTGCTATATCTTGCTTAAGAGAGAGAGGTCTTATGCGAGCACTCCTTTTTC 660
Db	
Qy	661 AATGAACCTCATCAGCAGAGATCAAGGACTTCACTGTGACTTTTGCTTGCCTGATGTTCCAA 720

Db 661 AATGAATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCTGCTGATTTCCAA 720
 QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCTCAA 780
 Db 721 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCTCAA 780
 QY 781 ATTGAGCAGAGATTTTAAACAGAAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTCATTT 840
 Db 781 ATTGAGCAGAGATTTTAAACAGAAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTCATTT 840
 QY 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGTTGGAATTCGATTTCTCA 900
 Db 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGTTGGAATTCGATTTCTCA 900
 QY 901 AAGGTTTTTCAGGCAGAAATCCTTTTGATTTTATGGAACACATTTCTTTAGAAGNAAA 960
 Db 901 AAGGTTTTTCAGGCAGAAATCCTTTTGATTTTATGGAACACATTTCTTTAGAAGNAAA 960
 QY 961 ACAAAATTTCTTTGAGAAAACGAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
 Db 961 ACAAAATTTCTTTGAGAAAACGAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
 QY 1021 ACAGATAACGTTCTTCACTTGGATGAGATTTT 1053
 Db 1021 ACAGATAACGTTCTTCACTTGGATGAGATTTT 1053

RESULT 3

BD093077
 LOCUS
 DEFINITION Novel protein and its DNA.
 ACCESSION BD093077
 VERSION BD093077.1 GI:22638665
 KEYWORDS WO 0100799-A/1.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1053)
 Nakamura, Y., Arakawa, H. and Tanaka, H.
 Novel protein and its DNA
 Patent: WO 0100799-A 1 04-JAN-2001;
 TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,
 HIROSHI TANAKA
 OS Homo sapiens (human)
 FN WO 0100799-A/1
 PD 04-JAN-2001
 PF 27-JUN-2000 WO 2000JP004189
 PR 28-JUN-1999 JP 99P 181131, 06-JUL-1999 JP 99P 192391 PR
 21-JAN-2000 JP 00P 017770
 PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
 C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC
 C12P21/02, A61K38/44,
 PC A61K46/00, C07K16/40, G01N33/53, C12Q1/26
 CC
 FH Key Location/Qualifiers.

FEATURES
 source
 1..1053
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 100.0%; Score 1053; DB 6; Length 1053;
 Best Local Similarity 100.0%; Pred. No. 1.2e-255;
 Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATCAGAGATCATTTCA 60
 Db 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATCAGAGATCATTTCA 60
 QY 61 GACACCAAGAGTGAATAAGTCAATGAAGCCACTCTCTAGAAGAGATTTCTCGC 120

Db 61 GACACCAAGAGTGAATAAGTCAATGAAGCCACTCTCTAGAAGAGATTTCTCGC 120
 QY 121 CGGTTTGTCTATCTTCCAAATCAGTACCTCATATTTGGAAATGTATAAAGCAGCAG 180
 Db 121 CGGTTTGTCTATCTTCCAAATCAGTACCTCATATTTGGAAATGTATAAAGCAGCAG 180
 QY 181 GCTTCTTCTGGACAGAGAGAGTCCAGTTCATCAAAAGGATCTCCCTCACCTGGAAACAAG 240
 Db 181 GCTTCTTCTGGACAGAGAGAGTCCAGTTCATCAAAAGGATCTCCCTCACCTGGAAACAAG 240
 QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTACGCTTTTGGCAGCCAGTGTAT 300
 Db 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTACGCTTTTGGCAGCCAGTGTAT 300
 QY 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTCAGGTTCCAGAGGCT 360
 Db 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTCAGGTTCCAGAGGCT 360
 QY 361 CGCTGTTTCTATGCTTTCAAAATTTCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 420
 Db 361 CGCTGTTTCTATGCTTTCAAAATTTCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 420
 QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATCAATTTGAA 480
 Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATCAATTTGAA 480
 QY 481 ACCATGCTTATGTAAGAAAACGAGATTTGGGCTTTGCGATGATGATAGAGATAAAA 540
 Db 481 ACCATGCTTATGTAAGAAAACGAGATTTGGGCTTTGCGATGATGATAGAGATAAAA 540
 QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGGAGTTTCTTCTCAGGA 600
 Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGGAGTTTCTTCTCAGGA 600
 QY 601 TCTTTTGTCTATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGAGTCACTTTTTC 660
 Db 601 TCTTTTGTCTATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGAGTCACTTTTTC 660
 QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGTGCTGATTTCCAA 720
 Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGTGCTGATTTCCAA 720
 QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
 Db 721 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
 QY 781 ATTGAGCAGAGATTTTAAACAGAAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTCATTT 840
 Db 781 ATTGAGCAGAGATTTTAAACAGAAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTCATTT 840
 QY 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGTTGGAATTCGATTTCTCA 900
 Db 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGTTGGAATTCGATTTCTCA 900
 QY 901 AAGGTTTTTCAGGCAGAAATCCTTTTGATTTTATGGAACACATTTCTTTAGAAGNAAA 960
 Db 901 AAGGTTTTTCAGGCAGAAATCCTTTTGATTTTATGGAACACATTTCTTTAGAAGNAAA 960
 QY 961 ACAAAATTTCTTTGAGAAAACGAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
 Db 961 ACAAAATTTCTTTGAGAAAACGAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
 QY 1021 ACAGATAACGTTCTTCACTTGGATGAGATTTT 1053
 Db 1021 ACAGATAACGTTCTTCACTTGGATGAGATTTT 1053

RESULT 4
 LOCUS
 DEFINITION Sequence 186 from Patent WO02068579.
 ACCESSION CQ714252
 CQ714252 1056 bp DNA linear PAT 03-FEB-2004
 CQ714252

```
VERSION      CQ714252.1  GI:42275109
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    1
AUTHORS      Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE        Kites, such as nucleic acid arrays, comprising a majority of
              humanexons or transcripts, for detecting expression and other uses
              thereof
JOURNAL      Patent: WO 02068579-A 186 06-SEP-2002;
              PE Corporation (NY) (US)
FEATURES     source
              1..1056
              /organism="Homo sapiens"
              /mol_type="unassigned DNA"
              /db_xref="taxon:9606"
ORIGIN
Query Match      100.0%; Score 1053; DB 6; Length 1056;
Best Local Similarity 100.0%; Pred. No. 1.2e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACCAACGAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 61 GACACCAACGAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 121 CGGTTGTGTCATCTTCCATCCAGTACCTGATATTGGAAATGTATAAACAGGCACAG 180
DB 121 CGGTTGTGTCATCTTCCATCCAGTACCTGATATTGGAAATGTATAAACAGGCACAG 180
QY 181 GCTTCCTCTCGACAGCAGAGGTCGACTTATCAAGGATCTCCTCACTGGAACAAG 240
DB 181 GCTTCCTCTCGACAGCAGAGGTCGACTTATCAAGGATCTCCTCACTGGAACAAG 240
QY 241 CTTAAAGCAGATGAGAAGTACTTCATCTCTCACATCTTAGCCCTTTTTCGAGCCAGTGAT 300
DB 241 CTTAAAGCAGATGAGAAGTACTTCATCTCTCACATCTTAGCCCTTTTTCGAGCCAGTGAT 300
QY 301 GGAATGTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCAGAGGCT 360
DB 301 GGAATGTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCAGAGGCT 360
QY 361 CGCTGTTCTATGGCTTCCAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG 420
DB 361 CGCTGTTCTATGGCTTCCAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
QY 481 ACCATGCCCTTGTGTTAAGAAAAAAGCAGATGGCCCTTGGATGGATAGAGATAGAAA 540
DB 481 ACCATGCCCTTGTGTTAAGAAAAAAGCAGATGGCCCTTGGATGGATAGAGATAGAAA 540
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTGTAAGAGAGTTTCTTCTCAGGA 600
DB 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTGTAAGAGAGTTTCTTCTCAGGA 600
QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGCTTTATGCCAGGACTCACTTTTTC 660
DB 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGCTTTATGCCAGGACTCACTTTTTC 660
QY 661 AATGAATCATCAGCAGAGATGAAGACATTCATGTGACATTTGCTTGCCTGATGTTCCAA 720
DB 661 AATGAATCATCAGCAGAGATGAAGACATTCATGTGACATTTGCTTGCCTGATGTTCCAA 720
QY 721 TACTTAGTAATAAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
DB 721 TACTTAGTAATAAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
QY 781 ATTGAGCAGAGTTTAAACAGAGCCTTCAGAGTTCGCTCATTTGGAATGAATTTGCAAT 840
DB 781 ATTGAGCAGAGTTTAAACAGAGCCTTCAGAGTTCGCTCATTTGGAATGAATTTGCAAT 840
QY 841 TTGATGAAACAGTACATTTGAGTTTGTAGCTCACAGATTTACTTTGTGGAACCTTGGATTTCA 900
DB 841 TTGATGAAACAGTACATTTGAGTTTGTAGCTCACAGATTTACTTTGTGGAACCTTGGATTTCA 900
QY 901 AAGTTTTTTCAGGAGAAAAATCCTTTTGTATTTTATGAAACATTTCTTTAGAGAAAA 960
DB 901 AAGTTTTTTCAGGAGAAAAATCCTTTTGTATTTTATGAAACATTTCTTTAGAGAAAA 960
QY 961 ACAAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTGTGAGTTATGCGCAGAAACC 1020
DB 961 ACAAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTGTGAGTTATGCGCAGAAACC 1020
QY 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053
DB 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053
RESULT 5
AB166671 1056 bp mRNA linear PRI 09-MAR-2004
LOCUS     Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase
DEFINITION small subunit 2, complete cds.
ACCESSION AB166671
VERSION    AB166671.1 GI:45259568
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1
AUTHORS    Ugai, H. and Yokoyama, K.K.
TITLE      Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase
JOURNAL    Published Only in Database (2004)
REFERENCE  2 (bases 1 to 1056)
AUTHORS    Ugai, H. and Yokoyama, K.K.
TITLE      Direct Submission
JOURNAL    Submitted (05-MAR-2004) Hideyo Ugai, RIKEN BioResource Center, Gene
              Engineering Division; 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074,
              Japan (E-mail:ugai@rtc.riken.go.jp, Tel:81-29-836-3612,
              Fax:81-29-836-9120)
FEATURES   Location/Qualifiers
              1..1056
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="p53R2.C19"
              /cell_line="Saos-2"
              /cell_type="Osteogenic sarcoma"
              1..1056
              /gene="p53R2"
              1..1056
              /gene="p53R2"
              /codon_start=1
              /product="p53-inducible ribonucleotide reductase small
              subunit 2"
              /protein_id="BAD12267.1"
              /db_xref="GI:45259569"
              /translation="MGDPERPEAAGLDQDERSSDPTNESEIKSNEEPLIKSSRRFVI
              PFIQYDPIWKMYKQAOASFTABEVDLSKDLPHNKLKXADEKYFISHILAFPAASDGI
              VNEULVERFSQEVQVPEARCFYQQLIENVHSEMYSLIIDTVISDPKKREFLFNALIE
              TMPYVKKADWALRWIADRSKTFGRVVFAFAAAGVGFSGSFAALFWLKKRGLMGLFT
              FSNELISRDEGLHCDPFLMFOYLVNKPSEERVREIIVDAVKIEQEFLEALPVGILG
              MNCILKQYIEFVADRLVLFGLFSGVSKVFAENPFDFMENISLEGKTNFPEKRVSEYORF
              AVMAETDNDVFTLDAF"
ORIGIN
```

Query Match	100.0%	Score 1053	DB 9	Length 1056
Best Local Similarity	100.0%	Pred. No. 1.2e-255		
Matches 1053	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	1	ATGGGCGACCCGGAAAGCGCGGAAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA	60	
Db	1	ATGGGCGACCCGGAAAGCGCGGAAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA	60	
Qy	61	GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC	120	
Db	61	GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC	120	
Qy	121	CGGTTTGTTCATCTTTTCCAAATCCAGTACCCCTGATATTTGGAAAAATGTATAAACAGGCACAG	180	
Db	121	CGGTTTGTTCATCTTTTCCAAATCCAGTACCCCTGATATTTGGAAAAATGTATAAACAGGCACAG	180	
Qy	181	GCTTCTCTTCGGACAGCAGAGAGAGGTGCACATTATCAAGAGATCTCCCTCACTGGAACAAG	240	
Db	181	GCTTCTCTTCGGACAGCAGAGAGAGGTGCACATTATCAAGAGATCTCCCTCACTGGAACAAG	240	
Qy	241	CTTAAAGCAGATGAGAAAGTACTTTCATCTCTCACATCTTAGCGCTTTTTCAGCAGCCAGTGAT	300	
Db	241	CTTAAAGCAGATGAGAAAGTACTTTCATCTCTCACATCTTAGCGCTTTTTCAGCAGCCAGTGAT	300	
Qy	301	GGAAATGTAAATGAATAAATTTGGTGGAGCGCTTTTATGTCAGGAGTGCAGGTTCCAGAGGCT	360	
Db	301	GGAAATGTAAATGAATAAATTTGGTGGAGCGCTTTTATGTCAGGAGTGCAGGTTCCAGAGGCT	360	
Qy	361	CGCTGTTTCTATGCGTTTCAAAATCTCATCGAGAATGTTCTACTCAGAGATGTACAGTTTG	420	
Db	361	CGCTGTTTCTATGCGTTTCAAAATCTCATCGAGAATGTTCTACTCAGAGATGTACAGTTTG	420	
Qy	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAAATTTTATTTTAAATGCAATTTGAA	480	
Db	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAAATTTTATTTTAAATGCAATTTGAA	480	
Qy	481	ACCATGCCCTATGTTTAAAGAAAAAGCAGATTTGGGCCCTTGGCGATGGATAGCAGATAGAAAA	540	
Db	481	ACCATGCCCTATGTTTAAAGAAAAAGCAGATTTGGGCCCTTGGCGATGGATAGCAGATAGAAAA	540	
Qy	541	TCTACTTTTGGGGAAAGAGTGGTGGCCCTTTGCTGCTGTAGAGAGGTTTCTTCTCAGGA	600	
Db	541	TCTACTTTTGGGGAAAGAGTGGTGGCCCTTTGCTGCTGTAGAGAGGTTTCTTCTCAGGA	600	
Qy	601	TCTTTTCTGCTATATTTCTGCTTAAAGAGAGAGGTCTTATGCCAGAGCTCACCTTTTTC	660	
Db	601	TCTTTTCTGCTATATTTCTGCTTAAAGAGAGAGGTCTTATGCCAGAGCTCACCTTTTTC	660	
Qy	661	AATGAACTCATCAGCAGAGATGAAGGACTTTCATCTGCTGCTTGGCTTGCCTGATGTTCCAA	720	
Db	661	AATGAACTCATCAGCAGAGATGAAGGACTTTCATCTGCTGCTTGGCTTGCCTGATGTTCCAA	720	
Qy	721	TACTTAGTAAATAAGCCTTCAGAAAGAGGTTCAGGAGATCATTTGTGTATGCTGTCAAA	780	
Db	721	TACTTAGTAAATAAGCCTTCAGAAAGAGGTTCAGGAGATCATTTGTGTATGCTGTCAAA	780	
Qy	781	ATTGAGCAGAGGTTTTTAAACAGAGCCCTGCCAGTTGGCCTCATTTGGAATGAAATGCAAT	840	
Db	781	ATTGAGCAGAGGTTTTTAAACAGAGCCCTGCCAGTTGGCCTCATTTGGAATGAAATGCAAT	840	
Qy	841	TTGATGAAACAGTACATTTGATTTGTAGCTTGACAGATTACTTTGTGGAACCTTGGATCTCA	900	
Db	841	TTGATGAAACAGTACATTTGATTTGTAGCTTGACAGATTACTTTGTGGAACCTTGGATCTCA	900	
Qy	901	AAGGTTTTTTCAGGCAGAAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGGAAAA	960	
Db	901	AAGGTTTTTTCAGGCAGAAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGGAAAA	960	
Qy	961	ACAAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCGAGTTATGGCAGAAACC	1020	
Db	961	ACAAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCGAGTTATGGCAGAAACC	1020	

Qy	1021	ACAGATAACGCTTCCACCTTGGATGCAGATTTT	1053
Db	1021	ACAGATACGCTTTCACCTTGGATGCAGATTTT	1053
RESULT 6			
AR454869			
LOCUS	AR454869	4955 bp	linear
DEFINITION	Sequence 3 from patent US 6682917.	DNA	PAT 20-FEB-2004
ACCESSION	AR454869		
VERSION	AR454869.1	GI:42688824	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unklassified.		
REFERENCE	1 (bases 1 to 4955)		
AUTHORS	Nakamura, Y., Arakawa, H. and Tanaka, H.		
TITLE	Protein having a ribonucleotide Reductase activity and a DNA thereof		
JOURNAL	Patent: US 6682917-A 3 27-JAN-2004;		
FEATURES	Location/Qualifiers		
source	1..4955		
	/organism="unknown"		
ORIGIN	/mol_type="genomic DNA"		
	Query Match	100.0%;	Score 1053; DB 6; Length 4955;
	Best Local Similarity	100.0%;	Pred. No. 1.1e-255;
	Matches 1053; Conservative	0; Mismatches	0; Indels 0; Gaps 0;
Qy	1	ATGGCGCACCCGGAAAGCCGGAAAGCCGGCGCTGGATCAGGATGAGAGATCATCTTCA	60
Db	245	ATGGCGCACCCGGAAAGCCGGAAAGCCGGCGCTGGATCAGGATGAGAGATCATCTTCA	304
Qy	61	GACACCAACGAAAGTGAAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC	120
Db	305	GACACCAACGAAAGTGAAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGAGTTCTCGC	364
Qy	121	CGGTTTGTGCATCTTTTCCAATCCAGTACCCCTGATATTTGGAAAAATGTATAAACAGGCACAG	180
Db	365	CGGTTTGTGCATCTTTTCCAATCCAGTACCCCTGATATTTGGAAAAATGTATAAACAGGCACAG	424
Qy	181	GCTTCTCTTTGGACAGCAGAGAGGTGCAGCTATATCAAAGGATCTCCCTCACTGGAAACAAG	240
Db	425	GCTTCTCTTTGGACAGCAGAGAGGTGCAGCTATATCAAAGGATCTCCCTCACTGGAAACAAG	484
Qy	241	CTTAAAGCAGATGAGAAAGTACTTCATCTCTCAACATCTAGCCCTTTTTCAGCCACGATGAT	300
Db	485	CTTAAAGCAGATGAGAAAGTACTTCATCTCTCAACATCTAGCCCTTTTTCAGCCACGATGAT	544
Qy	301	GGAAATCTAAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT	360
Db	545	GGAAATCTAAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT	604
Qy	361	CGCTGTGTTCTATGGCTTTTCAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG	420
Db	605	CGCTGTGTTCTATGGCTTTTCAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG	664
Qy	421	CTGATAGACACTTACATCAGAGATGCCAAGAAAGGGAAATTTTATTTAATGCAATTTGAA	480
Db	665	CTGATAGACACTTACATCAGAGATGCCAAGAAAGGGAAATTTTATTTAATGCAATTTGAA	724
Qy	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGGATGATAGCAGATAGAAAA	540
Db	725	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGGATGATAGCAGATAGAAAA	784
Qy	541	TCTACTTTTGGGAAAAAGATGGTGGCCCTTTGCTGTGTAGAAGGAGTTTCTTCTCAGGA	600
Db	785	TCTACTTTTGGGAAAAAGATGGTGGCCCTTTGCTGTGTAGAAGGAGTTTCTTCTCAGGA	844
Qy	601	TCTTTTCTGCTATATTCTGGCTAAAGACAGAGGCTTTATATGCCAGGACTCATCTTTTCC	660
Db	845	TCTTTTCTGCTATATTCTGGCTAAAGACAGAGGCTTTATATGCCAGGACTCATCTTTTCC	904

QY	661	AATGAACATCATCAGCAGAGATGAAGACTTCTACTGTGACTTTGCTTGCCTGATGTTCCAA	720
Db	905	AATGAACATCATCAGCAGAGATGAAGACTTCTACTGTGACTTTGCTTGCCTGATGTTCCAA	964
QY	721	TACTTAGTAATAAGCCTTCAGAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA	780
Db	965	TACTTAGTAATAAGCCTTCAGAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA	1024
QY	781	ATTGAGCAGAGTTTAAACAGAGCTTGCAGAGTTGGCCTCATTTGGAATGAATTCGATT	840
Db	1025	ATTGAGCAGAGTTTAAACAGAGCTTGCAGAGTTGGCCTCATTTGGAATGAATTCGATT	1084
QY	841	TTGATGAACAGTACATTCAGTTTGTAGCTGCAGAGTACTTGTGGAACCTTGGATTCTCA	900
Db	1085	TTGATGAACAGTACATTCAGTTTGTAGCTGCAGAGTACTTGTGGAACCTTGGATTCTCA	1144
QY	901	AAGGTTTTTTCAGGAGAAATCCTTTTGTATTTATGAGAAACATTTCTTTAGAGGAAA	960
Db	1145	AAGGTTTTTTCAGGAGAAATCCTTTTGTATTTATGAGAAACATTTCTTTAGAGGAAA	1204
QY	961	ACAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTTATGGCAGAAACC	1020
Db	1205	ACAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTTATGGCAGAAACC	1264
QY	1021	ACAGATAACGCTTCTACCTTGGATGCAGATTTT	1053
Db	1265	ACAGATAACGCTTCTACCTTGGATGCAGATTTT	1297
RESULT 7			
LOCUS	BD064765	4955 bp	DNA linear PAT 27-AUG-2002
DEFINITION	Novel protein and DNA thereof.		
ACCESSION	BD064765		
VERSION	BD064765.1	GI:22610368	
KEYWORDS	JP 2001269184-A/2.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 4955)		
TITLE	Nakamura, Y., Arakawa, H. and Tanaka, H.		
JOURNAL	Novel protein and DNA thereof		
COMMENT	Patent: JP 2001269184-A 2 02-OCT-2001; YUSUKE NAKAMURA, TAKEEDA CHEMICAL INDUSTRIES LTD OS Homo sapiens (human) PN JP 2001269184-A/2 PD 02-OCT-2001 PP 27-JUN-2000 JP 2000192401 PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC C12N15/09, A61K31/711, A61K38/00, A61K38/43, A61K45/00, A61K48/00, PC A61P35/00, PC A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC C12N5/10, C12N9/02, PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53/// (C12N15/09, PC C12R1:91), PC (C12N15/09, C12R1:19), (C12N9/02, C12R1:91), (C12P21/02, C12R1:91), PC (C12P21/08, C12R1:91), (C12N15/00, A61K37/02, A61K37/48, C12N5/00, PC (C12N15/00, C12R1:91), (C12N15/00, C12R1:19) CC Novel protein and DNA thereof FH Key Location/Qualifiers FT source 1. .4955 FT /organism= 'Homo sapiens (human)'. FEATURES source 1. .4955 Location/Qualifiers /organism= "Homo sapiens" /mol_type= "genomic DNA" /db_xref= "taxon:9606"		
ORIGIN	Query Match 100.0%; Score 1053; DB 6; Length 4955; Best Local Similarity 100.0%; Pred. No. 1.1e-255;		

Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 ATGGGGCAGCCGGAAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db	245 ATGGGGCAGCCGGAAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
Qy	61 GACACCAACGAAGTGAATTAAGTCAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 120
Db	305 GACACCAACGAAGTGAATTAAGTCAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 364
Qy	121 CGGTTTGTCACTTTCCAAATCCAGTACCTGATATTTGGAAAATGTTATAAACAGGCACAG 180
Db	365 CGGTTTGTCACTTTCCAAATCCAGTACCTGATATTTGGAAAATGTTATAAACAGGCACAG 424
Qy	181 GCCTTCCTTTCTGGACAGCAGAGAGGTGCACTTATCAAAGGATCTCCCTCACTGGAAACAAG 240
Db	425 GCCTTCCTTTCTGGACAGCAGAGAGGTGCACTTATCAAAGGATCTCCCTCACTGGAAACAAG 484
Qy	241 CTTAAAGCAGATGAGAAAGTACTTTCATCTCTCACATCTTTAGCCCTTTTTCGAGCCAGTGAT 300
Db	485 CTTAAAGCAGATGAGAAAGTACTTTCATCTCTCACATCTTTAGCCCTTTTTCGAGCCAGTGAT 544
Qy	301 GGAATTTGTAATGAAATTTGGTGAGCGCTTTAGTCAGAGGTGCAAGTTTCAGAGGCT 360
Db	545 GGAATTTGTAATGAAATTTGGTGAGCGCTTTAGTCAGAGGTGCAAGTTTCAGAGGCT 604
Qy	361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 420
Db	605 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 664
Qy	421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db	665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 724
Qy	481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGGAATAGCAGATAGAAA 540
Db	725 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGGAATAGCAGATAGAAA 784
Qy	541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
Db	785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 844
Qy	601 TCCTTTGCTGCTATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGAGCTCACTTTTTC 660
Db	845 TCCTTTGCTGCTATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGAGCTCACTTTTTC 904
Qy	661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db	905 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
Qy	721 TACTTAGTAATAAGCCTTCAGAGAAAGGTCAGGAGATCATTTGATGCTGTCAAA 780
Db	965 TACTTAGTAATAAGCCTTCAGAGAAAGGTCAGGAGATCATTTGATGCTGTCAAA 1024
Qy	781 ATTGAGCAGAGTTTAAACAGAGCTTGCAGAGTTGGCCTCATTTGGAATGAATTCGATT 840
Db	1025 ATTGAGCAGAGTTTAAACAGAGCTTGCAGAGTTGGCCTCATTTGGAATGAATTCGATT 1084
Qy	841 TTGATGAACAGTACATTCAGTTTGTAGCTGCAGAGTACTTGTGGAACCTTGATTTCTCA 900
Db	1085 TTGATGAACAGTACATTCAGTTTGTAGCTGCAGAGTACTTGTGGAACCTTGATTTCTCA 1144
Qy	901 AAGGTTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGGAACATTTCTTTAGAGGAAA 960
Db	1145 AAGGTTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGGAACATTTCTTTAGAGGAAA 1204
Qy	961 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGTTATGGCAGAAACC 1020
Db	1205 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGTTATGGCAGAAACC 1264
Qy	1021 ACAGATAACGCTTCTACCTTGGATGCAGATTTT 1053
Db	1265 ACAGATAACGCTTCTACCTTGGATGCAGATTTT 1297

Qy	481	ACCATGCCCTATGTTTAAGAAAAAAGCAGATTGGGCTTGGCATGGATAGCAGATAGAAAA	540
Db	725	ACCATGCCCTATGTTTAAGAAAAAAGCAGATTGGGCTTGGCATGGATAGCAGATAGAAAA	784
Qy	541	TCCTACTTTTGGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTCTTCTCTCAGGA	600
Db	785	TCCTACTTTTGGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTCTTCTCTCAGGA	844
Qy	601	TCCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCGCAGGACTCCTCTTTTCC	660
Db	845	TCCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCGCAGGACTCCTCTTTTCC	904
Qy	661	AATGAACCTCATGACGACAGATGAAGACCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA	720
Db	905	AATGAACCTCATGACGACAGATGAAGACCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA	964
Qy	721	TACTTAGTAAATAAGCCCTTCAGAAAGAAAGGTCAGGAGATCAATCTGTCATGCTGTCAAA	780
Db	965	TACTTAGTAAATAAGCCCTTCAGAAAGAAAGGTCAGGAGATCAATCTGTCATGCTGTCAAA	1024
Qy	781	ATTGAGCAGGAGTGTTTTAAACAGAAAGCTTGCCTTGGCTTCATTTGGAATGAATTCGATT	840
Db	1025	ATTGAGCAGGAGTGTTTTAAACAGAAAGCTTGCCTTGGCTTCATTTGGAATGAATTCGATT	1084
Qy	841	TTGATGAACAGTACATTCAGATTGTTAGCTGACAGAGATTAATCTTGTGAACTTGTCTCA	900
Db	1085	TTGATGAACAGTACATTCAGATTGTTAGCTGACAGAGATTAATCTTGTGAACTTGTCTCA	1144
Qy	901	AAGGTTTTTTCAGCAGAAATCTCTTTTGATTTTATGAAAAATCTTCTTTAGAAAGAAAA	960
Db	1145	AAGGTTTTTTCAGCAGAAATCTCTTTTGATTTTATGAAAAATCTTCTTTAGAAAGAAAA	1204
Qy	961	ACAAATTTCTTTGAGAAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC	1020
Db	1205	ACAAATTTCTTTGAGAAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC	1264
Qy	1021	ACAGATAACGCTCTTCACCTTGGATGACAGATTTT	1053
Db	1265	ACAGATAACGCTCTTCACCTTGGATGACAGATTTT	1297
RESULT 9			
LOCUS	AB036063	4955 bp mRNA linear	PRI 11-MAR-2000
DEFINITION	Homo sapiens p53R2 mRNA for ribonucleotide reductase, complete cds.		
ACCESSION	AB036063		
VERSION	AB036063.1	GI:7229085	
KEYWORDS	p53R2; ribonucleotide reductase.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Tanaka,H., Arakawa,H., Yamaguchi,T., Shiraishi,K., Fukuda,S., Matsui,K., Takei,Y. and Nakamura,Y.		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	A ribonucleotide reductase gene involved in a p53-dependent cell-cycle checkpoint for DNA damage		
JOURNAL	Nature	404 (6773), 42-49 (2000)	
MEDLINE	20179179		
PUBMED	10716435		
REFERENCE	2 (bases 1 to 4955)		
AUTHORS	Tanaka,H., Arakawa,H. and Nakamura,Y.		
TITLE	Submitted (18-DEC-1999) Yusuke Nakamura, University of Tokyo, Institute of Medical Science, Human Genome Center, Laboratory of Molecular Medicine; 4-6-1 Shiokanadai, Minato-ku, Tokyo 108-8639, Japan [E-mail:yusuke@ims.u-tokyo.ac.jp, Tel:+81-3-5449-5372, Fax:+81-3-5449-5433]		
JOURNAL			
FEATURES	Location/Qualifiers		
source	1. .4955		
	/organism="Homo sapiens"		
	/mol type="mRNA"		

/db_xref="taxon:9606"
/chromosome="8"
/map="8q23.1"
1..4955
/gene="p53R2"
245..1300
/gene="p53R2"
/codon_start=1
/product="ribonucleotide reductase"
/protein_id="BA092434.1"
/db_xref="GI:7229086"
/translation="MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSRFVI
PFIQPDIMKMTQKQASFWTAEEVDLSKDLPHWNKLKADKRYFISHILAFPAASDGI
VNIENLVKFSQVQPEARCFYGOILINHVSEMYSLLDIVIRDPKKRFLFNALIE
TMPVYKADWALRIADRKSTFGERVVAFAAEGVFPSGFAAIFMLKRLGLMPLGT
FSNELSRDGLHCFACLMFQYLUNKPSEERVELIIVDAVKIEQBFLEALPVLGLG
MNCILMKQYIEFVADRLILVELGFSKVFQAEENPFDFMENISLSEKTNFFEKRVSEYORF
AVMAETDNDVFTLDADF"
4933
/gene="p53R2"
/note="22 a nucleotides"

polya_site

ORIGIN

Query Match 100.0%; Score 1053; DB 9; Length 4955;
Best Local Similarity 100.0%; Pred. No. 1.1e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTTCAC 964
Qy 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 965 TACTTAGTAAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 1024
Qy 781 ATTGAGCAGGAGTTTTTAACAGAAAGCCTTGCCAGTTTGGCCCTCATTTGGAATGAATTCGATT 840
Db 1025 ATTGAGCAGGAGTTTTTAACAGAAAGCCTTGCCAGTTTGGCCCTCATTTGGAATGAATTCGATT 1084
Qy 841 TTGATGAACAGTACATGATGATTTGTAGCTGACAGATTAATTTGTTGGAACCTTGGATCTCA 900
Db 1085 TTGATGAACAGTACATGATGATTTGTAGCTGACAGATTAATTTGTTGGAACCTTGGATCTCA 1144
Qy 901 AAGGTTTTTCAGGACAGAAATCCTTTTGTATTTATGAAACATTTCTTTTAAAGGAAAA 960
Db 1145 AAGGTTTTTCAGGACAGAAATCCTTTTGTATTTATGAAACATTTCTTTTAAAGGAAAA 1204
Qy 961 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGGTTTATGCGAGAAACC 1020
Db 1205 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGGTTTATGCGAGAAACC 1264
Qy 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
Db 1265 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1297

RESULT 10
AR454877
LOCUS
DEFINITION AR454877 1053 bp DNA linear PAT 20-FEB-2004
ACCESSION AR454877
VERSION AR454877.1 GI:42688832
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1053)
AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE Protein having a ribonucleotide Reductase activity and a DNA
JOURNAL Patent: US 6682917-A 12 27-JAN-2004;
FEATURES
source
1..1053
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 99.8%; Score 1051.4; DB 6; Length 1053;
Best Local Similarity 99.9%; Pred. No. 3e-255;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGCGGACCCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
Db 1 ATGGCGGACCCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
Qy 61 GACACCAAGAAAGTGAATAAAGTCAAAATGAAGAGCACTCCTTAAGAAAGATTCCTGC 120
Db 61 GACACCAAGAAAGTGAATAAAGTCAAAATGAAGAGCACTCCTTAAGAAAGATTCCTGC 120
Qy 121 CGGTTTGTCACTTTTCCAATCCAGTACCTGATATTTGGAAAAATGTATAAAGGACAG 180
Db 121 CGGTTTGTCACTTTTCCAATCCAGTACCTGATATTTGGAAAAATGTATAAAGGACAG 180
Qy 181 GCTTCTCTTCTGGACAGCAGAGAGTGGCTTATCAAGAGATCTCCCTCACTGGAAACAG 240
Db 181 GCTTCTCTTCTGGACAGCAGAGAGTGGCTTATCAAGAGATCTCCCTCACTGGAAACAG 240
Qy 241 CTTTAAAGCAGATGAGAGTACTTCACTCTCAATCTTAGCCTTTTTCGACCCAGTAT 300
Db 241 CTTTAAAGCAGATGAGAGTACTTCACTCTCAATCTTAGCCTTTTTCGACCCAGTAT 300
Qy 301 GGAATGTAAATGNAATTTGGTGGAGCGCTTTTAGTCAGGAGTTCAGAGGCT 360

```
Db 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
Qy 361 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAGAATGTTTCCACATCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAGAATGTTTCCACATCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAAAA 540
Qy 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
Qy 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCGAGGACTCACTTTTTC 660
Db 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCGAGGACTCACTTTTTC 660
Qy 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGGCTTGGCTGATGTTCCAA 720
Db 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGGCTTGGCTGATGTTCCAA 720
Qy 721 TACTTAGTAAATAGGCTTCAAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 721 TACTTAGTAAATAGGCTTCAAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAGGCTTGCAGAGTGGCTCATTTGGAATGAATTCGATT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGGCTTGCAGAGTGGCTCATTTGGAATGAATTCGATT 840
Qy 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTGTGAACTTTGATTTCTCA 900
Db 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTGTGAACTTTGATTTCTCA 900
Qy 901 AAGGTTTTTTCAGGAGAAATCCTTTGATTTTATGGAAGAAATTTCTTTAGAGGAA 960
Db 901 AAGGTTTTTTCAGGAGAAATCCTTTGATTTTATGGAAGAAATTTCTTTAGAGGAA 960
Qy 961 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGGTTTTCAGGTTTATGCGAGAAACC 1020
Db 961 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGGTTTTCAGGTTTATGCGAGAAACC 1020
Qy 1021 ACAGATAAGCTCTTACCTTTGGATGAGATTTT 1053
Db 1021 ACAGATAAGCTCTTACCTTTGGATGAGATTTT 1053

RESULT 11
LOCUS BD064773 1053 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel protein and DNA thereof.
ACCESSION BD064773
VERSION BD064773.1 GI:22610376
KEYWORDS JP 2001269184-A/10.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1053)
AUTHORS Nakamura, Y., Arakawa, H. and Tanaka, H.
TITLE Novel protein and DNA thereof
JOURNAL Patent: JP 2001269184-A 10 02-OCT-2001;
YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD
COMMENT OS Homo sapiens (human)
PN JP 2001269184-A/10
PD 02-OCT-2001
PF 27-JUN-2000 JP 2000192401
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
```

```
FEATURES
source 1..1053
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 1053;
Best Local Similarity 99.9%; Pred. No. 3e-255;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGCGGACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGCGGACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Qy 61 GACACCAAGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTGC 120
Db 61 GACACCAAGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTGC 120
Qy 121 CGTTTGTGTCATCTTCCATCCAGTACCTGATATTTGGAATAATGTAACAGGACAG 180
Db 121 CGTTTGTGTCATCTTCCATCCAGTACCTGATATTTGGAATAATGTAACAGGACAG 180
Qy 181 GCTTCTCTTGGACAGCAGAGAGGTGACATTTATCAAGAGATCTCCCTCAGTGGAAACAG 240
Db 181 GCTTCTCTTGGACAGCAGAGAGGTGACATTTATCAAGAGATCTCCCTCAGTGGAAACAG 240
Qy 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGTAT 300
Db 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGTAT 300
Qy 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGCAGGTTCCAGAGGT 360
Db 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGCAGGTTCCAGAGGT 360
Qy 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAAAA 540
Qy 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
Qy 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGGCTTTATGCGAGGACTCACTTTTTC 660
Db 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCGAGGACTCACTTTTTC 660
Qy 661 AATGAACATCATCAGCAGAGATGAAGAGTTCACCTGTGACTTTTGTGCTGATGTTCCAA 720
Db 661 AATGAACATCATCAGCAGAGATGAAGAGTTCACCTGTGACTTTTGTGCTGATGTTCCAA 720
Qy 721 TACTTAGTAAATAGGCTTCAAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 780
```

Db 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTCATT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTCATT 840
Qy 841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGAACTTTGGATTCTCA 900
Db 841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGAACTTTGGATTCTCA 900
Qy 901 AAGGTTTTTTCAGCAGAAAATCCTTTTGAATTTATGAAAACATTTCTTTAGAGGAAAA 960
Db 901 AAGGTTTTTTCAGCAGAAAATCCTTTTGAATTTATGAAAACATTTCTTTAGAGGAAAA 960
Qy 961 ACAAAATTTCTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGGATGAGCAAAACC 1020
Db 961 ACAAAATTTCTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGGATGAGCAAAACC 1020
Qy 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053

RESULT 12
BD093086
LOCUS
DEFINITION
Novel protein and its DNA.
ACCESSION
BD093086
VERSION
BD093086.1 GI:22638674
KEYWORDS
WO 0100799-A/10.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1053)
AUTHORS
Nakamura, Y., Arakawa, H. and Tanaka, H.
TITLE
Novel protein and its DNA
JOURNAL
TAKEDA: CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,
HIROSHI TANAKA
COMMENT
OS Homo sapiens (human)
PN WO 0100799-A/10
PD 04-JAN-2001
PF 27-JUN-2000 WO 2000JP004189
PR 28-JUN-1999 JP 99P 181131, 06-JUL-1999 JP 99P 192391 PR
21-JAN-2000 JP 00P 017770
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC
C12P21/02, A61K38/44,
PC A61K48/00, C07K16/40, G01N33/53, C12Q1/26
CC
FH Key Location/Qualifiers.
FEATURES
source
1..1053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

Query Match 99.8%; Score 1051.4; DB 6; Length 1053;
Best Local Similarity 99.8%; Pred. No. 3e-255;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGGGGGACCGGAAAGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGGGGACCGGAAAGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Qy 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Qy 121 CGGTTTGTATCTTTCAATCCAGTACCTCGATATTGGAATGATATAACAGGCACAG 180

Db 121 CGGTTTGTATCTTTCAATCCAGTACCTCGATATTGGAATGATATAACAGGCACAG 180
Qy 181 GCTTCCTTCTGGACAGCAGAGAGGTGACTTATCAAGAGGATCTCCCTCACTGGACAG 240
Db 181 GCTTCCTTCTGGACAGCAGAGAGGTGACTTATCAAGAGGATCTCCCTCACTGGACAG 240
Qy 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGGCCAGTGT 300
Db 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGGCCAGTGT 300
Qy 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAAGAGGTGCGAGTTCAGAGGCT 360
Db 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAAGAGGTGCGAGTTCAGAGGCT 360
Qy 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGATGTTTCACTCAGAGATGTACAGTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGATGTTTCACTCAGAGATGTACAGTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAAATGCAATTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAAATGCAATTGAA 480
Qy 481 ACCATGCCCTATGTTAAAGAAAAAGCAGATTGGCCCTTTGCGATGATAGCAGATGAGAAA 540
Db 481 ACCATGCCCTATGTTAAAGAAAAAGCAGATTGGCCCTTTGCGATGATAGCAGATGAGAAA 540
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTCTGCTGCTAGAGAGGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTCTGCTGCTAGAGAGGTTTCTTCTCAGGA 600
Qy 601 TCTTTTCTGCTATATTCTGCTAAAGAGAGAGTCTTATGCCAGGAGTCACTTTTCTCC 660
Db 601 TCTTTTCTGCTATATTCTGCTAAAGAGAGAGTCTTATGCCAGGAGTCACTTTTCTCC 660
Qy 661 AATGAACCTCATCAGCAGAGATGAAGAGCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGAGCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Qy 721 TACTTAGTAAATAGCCTTCAGAAAGAGGCTGAGGAGATCATTTGTGATGCTGTCAAA 780
Db 721 TACTTAGTAAATAGCCTTCAGAAAGAGGCTGAGGAGATCATTTGTGATGCTGTCAAA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTGGCCTCATTTGGAATGAATTCATT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTGGCCTCATTTGGAATGAATTCATT 840
Qy 841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Db 841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Qy 901 AAGGTTTTTTCAGCAGAAAATCCTTTTGAATTTATGAAAACATTTCTTTAGAGGAAAA 960
Db 901 AAGGTTTTTTCAGCAGAAAATCCTTTTGAATTTATGAAAACATTTCTTTAGAGGAAAA 960
Qy 961 ACAAAATTTCTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGGATGAGCAAAACC 1020
Db 961 ACAAAATTTCTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGGATGAGCAAAACC 1020
Qy 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053

RESULT 13
AR454870
LOCUS
DEFINITION
Sequence 4 from patent US 6682917.
ACCESSION
AR454870
VERSION
AR454870.1 GI:42688825
KEYWORDS
Unknown.
ORGANISM
Unknown.

REFERENCE	Unclassified.									
AUTHORS	1 (bases 1 to 1081)									
TITLE	Nakamura,Y., Arakawa,H. and Tanaka,H.									
FEATURES	Protein having a ribonucleotide Reductase activity and a DNA thereof									
source	Patent: US 6682917-A 4 27-JAN-2004;									
ORIGIN	Location/Qualifiers									
	1..1081									
	/organism="unknown"									
	/mol_type="genomic DNA"									
	Query Match	99.8%;	Score 1051.4;	DB 6;	Length 1081;					
	Best Local Similarity	99.9%;	Pred. No. 3e-255;							
	Matches 1052;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;					
Qy	1	ATGGGGCACCCGGAAAGCCGGAGAGCGCGCGCTGGATCAGGATCAGGATGAGAGATCATCTTCA	60							
Db	20	ATGGGCGACCCGGAAAGCCGGAGAGCGCGCTGGATCAGGATGAGAGATCATCTTCA	79							
Qy	61	GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTGCG	120							
Db	80	GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTGCG	139							
Qy	121	CGGTTTCTCATCTTTCCAAATCCAGTACCTCGATATTTGGAAAAATGTATAAACAGGCACAG	180							
Db	140	CGGTTTCTCATCTTTCCAAATCCAGTACCTCGATATTTGGAAAAATGTATAAACAGGCACAG	199							
Qy	181	GCTTCTCTTCTGGACAGCAGAGAGGTCGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG	240							
Db	200	GCTTCTCTTCTGGACAGCAGAGAGGTTGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG	259							
Qy	241	CTTAAAGCAGATGAGAAAGTACTTCATCTCTCACATCTTTAGGCGCTTTTTCGACGACGTGAT	300							
Db	260	CTTAAAGCAGATGAGAAAGTACTTCATCTCTCACATCTTTAGGCGCTTTTTCGACGACGTGAT	319							
Qy	301	GGAAATCTGAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGTCAGGTTCCAGAGGCT	360							
Db	320	GGAAATCTGAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGTCAGGTTCCAGAGGCT	379							
Qy	361	CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAGAAATTTCACTCAGAGATGTACAGTTTG	420							
Db	380	CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAGAAATTTCACTCAGAGATGTACAGTTTG	439							
Qy	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTTGAA	480							
Db	440	CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTTGAA	499							
Qy	481	ACCATGCCCTATGTTAAGAAAAGCAGATTGGGCGCTTGGATGGATAGCAGATAGAAAA	540							
Db	500	ACCATGCCCTATGTTAAGAAAAGCAGATTGGGCGCTTGGATGGATAGCAGATAGAAAA	559							
Qy	541	TCTACTTTTGGGGAAGAGTGGTGGCGCTTTGCTGCTGTAGAAGAGATTTTCTCTCAGGA	600							
Db	560	TCTACTTTTGGGGAAGAGTGGTGGCGCTTTGCTGCTGTAGAAGAGATTTTCTCTCAGGA	619							
Qy	601	TCTTTTGCTGCTATATTTCTGCTTAAAGAGAGAGTCTTATGCCAGAGCTACCTTTTTTCC	660							
Db	620	TCTTTTGCTGCTATATTTCTGCTTAAAGAGAGAGTCTTATGCCAGAGCTACCTTTTTTCC	679							
Qy	661	AATGAACTCATCAGCAGAGATGAAGGACTTCACCTGTGACTTTGCTGCTGATGTTCCAA	720							
Db	680	AATGAACTCATCAGCAGAGATGAAGGACTTCACCTGTGACTTTGCTGCTGATGTTCCAA	739							
Qy	721	TACTTAGTAAATAGCCTTCAGAAGAAAAGGTTCAGGAGATCATTTGTGATGCTGTCAAA	780							
Db	740	TACTTAGTAAATAGCCTTCAGAAGAAAAGGTTCAGGAGATCATTTGTGATGCTGTCAAA	799							
Qy	781	ATTGAGCAGAGTTTTTAAACAGAGGCTTCCCGAGTTGGCGCTCATTTGGAAATGAAATTCATT	840							
Db	800	ATTGAGCAGAGTTTTTAAACAGAGGCTTCCCGAGTTGGCGCTCATTTGGAAATGAAATTCATT	859							
Qy	841	TTGATGAAAACAGTACATTTGATGTTGTAGCTGACAGATTACTTTGTGGAACTTGGATTTCTCA	900							

```
QY 181 GCTTCTCTCTGACAGCAGAGAGGTCGACTTATCAAGAGTCTCCCTCACTGGACACAG 240
Db 200 GCTTCTCTCTGACAGCAGAGAGGTTGACTTATCAAGAGTCTCCCTCACTGGACACAG 259
QY 241 CTTAAAGCAGATGAGAGTACTTCACTCTCTCACTCTAGCCCTTTTTCGAGCCAGTGAT 300
Db 260 CTTAAAGCAGATGAGAGTACTTCACTCTCTCACTCTAGCCCTTTTTCGAGCCAGTGAT 319
QY 301 CGAAATGTAATGAAATTTTGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Db 320 GGAATGTAATGAAATTTTGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 379
QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTATCAGAGATGTTCACTCAGAGATGACAGTTTG 420
Db 380 CGCTGTTTCTATGGCTTTCAAATTTCTATCAGAGATGTTCACTCAGAGATGACAGTTTG 439
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 499
QY 481 ACCATGCCCTATGTTAAGAAAGGAGAGATGGGCGCTTGGATGGATAGCAGATAGAAA 540
Db 500 ACCATGCCCTATGTTAAGAAAGGAGAGATGGGCGCTTGGATGGATAGCAGATAGAAA 559
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGCTGTAGAGAGTCTTCTCTCAGGA 600
Db 560 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGCTGTAGAGAGTCTTCTCTCAGGA 619
QY 601 TCTTTTGTGCTATATCTGGCTTAAAGAGAGAGGTCTTATGCGCAGGACTCACTTTTTC 660
Db 620 TCTTTTGTGCTATATCTGGCTTAAAGAGAGAGGTCTTATGCGCAGGACTCACTTTTTC 679
QY 661 AATGAACACTCAGCAGAGATGAGAGATCTCACTGTGACTTTGCTTGGCTGATGTTCCAA 720
Db 680 AATGAACACTCAGCAGAGATGAGAGATCTCACTGTGACTTTGCTTGGCTGATGTTCCAA 739
QY 721 TACTTAGTAAATAAGCTTTTAAAGAGAGAGGTGAGAGATCACTGTGACTTTGCTTGGCTGATGTTCCAA 780
Db 740 TACTTAGTAAATAAGCTTTTAAAGAGAGAGGTGAGAGATCACTGTGACTTTGCTTGGCTGATGTTCCAA 799
QY 781 ATTGAGCAGGAGTTTAAAGAGAGCTTTGCCAGTTGGCTCATTTGGAATGCAATT 840
Db 800 ATTGAGCAGGAGTTTAAAGAGAGCTTTGCCAGTTGGCTCATTTGGAATGCAATT 859
QY 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAAGTGTGAACTTGAATCTCA 900
Db 860 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAAGTGTGAACTTGAATCTCA 919
QY 901 AAGGTTTTTTCAGCAGAGAAATCCTTTTGTATTTTATGAAACATTTCTTTAGAGGAAA 960
Db 920 AAGGTTTTTTCAGCAGAGAAATCCTTTTGTATTTTATGAAACATTTCTTTAGAGGAAA 979
QY 961 ACAAAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
Db 980 ACAAAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1039
QY 1021 ACAGATAACGCTCTTCACTTTGGATGAGATTTT 1053
Db 1040 ACAGATAACGCTCTTCACTTTGGATGAGATTTT 1072
```

RESULT 15

```
BD093079
LOCUS BD093079 1081 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel protein and its DNA.
ACCESSION BD093079
VERSION BD093079.1 GI:22638667
KEYWORDS WO 0100799-A/3.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1081)
```

AUTHORS
TITLE
JOURNAL

Nakamura, Y., Arakawa, H. and Tanaka, H.
Novel protein and its DNA
Patent: WO 0100799-A 3 04-JAN-2001;
TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,
HIROSHI TANAKA

COMMENT

OS Homo sapiens (human)
PN WO 0100799-A/3
PD 04-JAN-2001
PF 27-JUN-2000 WO 2000JP004189
PR 28-JUN-1999 JP 99P 181131, 06-JUL-1999 JP 99P 192391 PR
21-JAN-2000 JP 00P 017770
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N5/04, C12N15/53, C12N1/15, C12N1/19, C12N5/10 PC
C12P21/02, A61K38/44,
PC A61K48/00, C07K16/40, G01N33/53, C12Q1/26
CC

FEATURES
source

Key Location/Qualifiers.
1. .1081
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 99.8%; Score 1051.4; DB 6; Length 1081;
Best Local Similarity 99.9%; Pred. No. 3e-255;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGGAAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 20 ATGGGCGACCCGGGAAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 79
QY 61 GACACCAACCAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 80 GACACCAACCAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 139
QY 121 CGGTTTGTCTATCTTCCAAATCCAGTACCTCTGATATTTGGAAAATGTATAACACAGCAG 180
Db 140 CGGTTTGTCTATCTTCCAAATCCAGTACCTCTGATATTTGGAAAATGTATAACACAGCAG 199
QY 181 GCTTCTCTTGGACAGCAGAGAGGTGACATTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db 200 GCTTCTCTTGGACAGCAGAGAGGTGACATTATCAAGAGATCTCCCTCACTGGAACAAG 259
QY 241 CTTAAAGCAGATGAGAGTACTTCACTCTCACTATCTAGCCCTTTTTCGAGCCAGTGAT 300
Db 260 CTTAAAGCAGATGAGAGTACTTCACTCTCACTATCTAGCCCTTTTTCGAGCCAGTGAT 319
QY 301 GGAATTTGTAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGAGGTTCCAGAGGCT 360
Db 320 GGAATTTGTAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGAGGTTCCAGAGGCT 379
QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAAGATGTTCTCTCAGAGATGTACAGTTTG 420
Db 380 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAAGATGTTCTCTCAGAGATGTACAGTTTG 439
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 499
QY 481 ACCATGCCCTATGTTAAGAAAGGAGAGATGGGCGCTTGGATGGATAGCAGATAGAAA 540
Db 500 ACCATGCCCTATGTTAAGAAAGGAGAGATGGGCGCTTGGATGGATAGCAGATAGAAA 559
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGCTGTAGAGAGTCTTCTCTCAGGA 600
Db 560 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGCTGTAGAGAGTCTTCTCTCAGGA 619
QY 601 TCTTTTGTGCTATATCTGGCTTAAAGAGAGAGGTCTTATGCGCAGGACTCACTTTTTC 660
Db 620 TCTTTTGTGCTATATCTGGCTTAAAGAGAGAGGTCTTATGCGCAGGACTCACTTTTTC 679
QY 661 AATGAACACTCAGCAGAGATGAGAGATCTCACTGTGACTTTGCTTGGCTGATGTTCCAA 720

```
Db      ||||||| 680 AATGAACTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 739
Qy      ||||||| 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGTCAGGGAGATCATTTGTTGATGCTGTCAAA 780
Db      ||||||| 740 TACTTAGTAAATAAGCCTTCAGAGAAAGGTCAGGGAGATCATTTGTTGATGCTGTCAAA 799
Qy      ||||||| 781 ATTGACGAGGAGTTTTTAACAGAAAGCCTTGCCAGTTGGCCCTCATTTGGAATGAATTGCATT 840
Db      ||||||| 800 ATTGACGAGGAGTTTTTAACAGAAAGCCTTGCCAGTTGGCCCTCATTTGGAATGAATTGCATT 859
Qy      ||||||| 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Db      ||||||| 860 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 919
Qy      ||||||| 901 AAGGTTTTTCAGGCAGAAATCCTTTTGTATTTTATGMAAACATTTCTTTAGAGGAAAA 960
Db      ||||||| 920 AAGGTTTTTCAGGCAGAAATCCTTTTGTATTTTATGMAAACATTTCTTTAGAGGAAAA 979
Qy      ||||||| 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTTATGGCAGAAACC 1020
Db      ||||||| 980 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTTATGGCAGAAACC 1039
Qy      ||||||| 1021 ACAGATAACGTTCTTACCTTGGATGAGATTTT 1053
Db      ||||||| 1040 ACAGATAACGTTCTTACCTTGGATGAGATTTT 1072
```

Search completed: October 30, 2005, 01:51:28
Job time : 3263 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2005, 21:38:30 ; Search time 440.5 Seconds
(without alignments)
14150.937 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1053

Sequence: 1 atggcgaccgcaagggcc.....tcacctgtgacgagatttt 1053

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1053	100.0	1053	4	Aaf32438 Human rib
2	1053	100.0	2596	3	Aa12411 CDNA enco
3	1053	100.0	4955	4	Aaf32439 Human rib
4	1053	100.0	4955	13	Adr24210 Breast ca
5	1051.4	99.8	1053	4	Aaf32447 Human rib
6	1051.4	99.8	1081	4	Aaf32440 Human rib
7	1049.8	99.7	1601	4	Aa14924 Human cDN
8	650.6	61.8	706	13	Ad57092 Novel can
9	578.4	54.9	1989	3	Aac78111 Human can
10	578.4	54.9	2215	12	Adk70302 Respirato
11	578.4	54.9	2216	10	Adj56536 Murine cD
12	578.4	54.9	2482	4	Aas44917 Human con
13	578.4	54.9	2500	6	Ab165414 Lung can
14	578.4	54.9	2500	6	Ab166517 Lung can
15	578.4	54.9	2500	6	Ab165859 Lung can
16	578.4	54.9	2500	8	Abx10335 DNA enco
17	578.4	54.9	2500	11	Adi32044 Human cDN
18	578.4	54.9	2500	12	Adn04443 Antipsoi
19	578.4	54.9	2500	12	Adq09273 Human RRM
20	578.4	54.9	2500	13	Acn37637 Tumour-as

ALIGNMENTS

RESULT 1

AAF32438
ID AAF32438 standard; cDNA; 1053 BP.

XX AAF32438;

DT 18-APR-2001 (first entry)

XX Human ribonucleotide reductase TP53R2H nucleotide sequence SEQ ID NO:2.

XX Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.

XX Homo sapiens.

XX WO200100799-A1.

XX 04-JAN-2001.

XX 27-JUN-2000; 2000WO-JF004189.

XX 28-JUN-1999; 99JP-00181131.

XX 06-JUL-1999; 99JP-00192391.

XX 21-JAN-2000; 2000JP-00017770.

XX (TAKE) TAKEDA CHEM IND LTD.

XX (NAKA/) NAKAMURA Y.

XX Nakamura Y, Arakawa H, Tanaka H;

XX WPI; 2001-112446/12.

XX P-PSDB; AAB69050.

XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for diagnosis, treatment and prevention of cancer.

XX Claim 5; Fig 1-3; 102pp; Japanese.

XX The present sequence encodes a human ribonucleotide reductase designated TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of the DNA repair mechanism and its activity is induced by p53. It can be used for the treatment, prevention and diagnosis of a wide range of cancers

XX Sequence 1053 BP; 314 A; 192 C; 246 G; 301 T; 0 U; 0 Other;

Aah73225 Human cer
Adn03788 Antipsoi
Ado19225 Human PRO
Aa44745 Human ful
Aav05641 Human rib
Aas41006 CDNA enco
Ab190228 Human pol
Aa810707 Human cDN
Aas79474 DNA enco
Ab114627 Drosophil
Aa62387 CDNA sequ
Aax51874 Human sec
Abt20895 Aspergill
Aa84544 Swinepox
Aa84949 Reverse c
Aa84948 Nucleotid
Ab114626 Drosophil
Aa858378 Bacteri
Abt19075 Aspergill
Adt47814 Bacteri
Ab322250 Candida a
Aa849379 Bacteri
Acc61238 Gene sequ
Adk63241 Disease t

21 578.4 54.9 3294 4 AAH73225
22 578.4 54.9 3393 12 ADN03788
23 578.4 54.9 3393 12 ADO19225
24 569 54.0 2641 4 AAS44745
25 568.8 54.0 1170 2 AAV05641
26 561 53.3 1328 12 ADO57308
27 547.8 52.0 978 4 AAS41006
28 545.4 51.8 977 6 ABL90228
29 528.8 50.2 608 4 AAO70707
30 514 48.8 1371 5 AAS79474
31 463 44.0 1289 4 ABL14627
32 442 42.0 1146 6 AAS62387
33 441 41.9 481 2 AAX51874
34 432.2 41.0 1218 8 AAT20895
35 429.8 40.8 14176 2 AAT84564
36 429.8 40.8 14176 4 AAF84949
37 429.8 40.8 14176 4 AAF84948
38 420.4 39.9 3945 4 ABL14626
39 403.2 38.3 963 13 ADS58378
40 397 37.7 1146 8 AAT19075
41 392.6 37.3 1173 13 ADT47814
42 392 37.2 1242 6 ABZ32250
43 386.6 36.7 1292 13 ADS49379
44 380.4 36.1 1200 10 ACC61238
45 380.4 36.1 1200 10 ADK63241


```
Db 56 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 115
Qy 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 116 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 175
Qy 121 CGGTTTGTCTATCTTCCAAATCCAGTACCCCTGATATTTGGAAATGTATAAACAGGCACAG 180
Db 176 CGGTTTGTCTATCTTCCAAATCCAGTACCCCTGATATTTGGAAATGTATAAACAGGCACAG 235
Qy 181 GCTTCCTTCTGGACAGCAGAAAGGTGCACTTATCAAGAGATCTCCCTCACTGGAAACAAG 240
Db 236 GCTTCCTTCTGGACAGCAGAAAGGTGCACTTATCAAGAGATCTCCCTCACTGGAAACAAG 295
Qy 241 CTTAAAGCAGATCAGAAAGTACTTCACTCTCTCACTCTAGCCCTTTTTCAGCCAGTGAT 300
Db 296 CTTAAAGCAGATCAGAAAGTACTTCACTCTCTCACTCTAGCCCTTTTTCAGCCAGTGAT 355
Qy 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAAGTTCCAGAGGCT 360
Db 356 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAAGTTCCAGAGGCT 415
Qy 361 CGCTGTTTCTATGCGCTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db 416 CGCTGTTTCTATGCGCTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 475
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 480
Db 476 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 535
Qy 481 ACCATGCCCTATGTTAAGAAAGAGCAGATTTGGGCTTGGATGGATAGCAGATAGAA 540
Db 536 ACCATGCCCTATGTTAAGAAAGAGCAGATTTGGGCTTGGATGGATAGCAGATAGAA 595
Qy 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAAGAGTTTCTTCTCAGGA 600
Db 596 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAAGAGTTTCTTCTCAGGA 655
Qy 601 TCTTTTGTCTATATCTGGCTTAAAGAGAGAGGCTTATGCAAGACTCACTTTTTC 660
Db 656 TCTTTTGTCTATATCTGGCTTAAAGAGAGAGGCTTATGCAAGACTCACTTTTTC 715
Qy 661 AATGAATCATCAGCAGAGATGAAGACTTCACTGTGACTTTTCTTGCCTGATGTTCCAA 720
Db 716 AATGAATCATCAGCAGAGATGAAGACTTCACTGTGACTTTTCTTGCCTGATGTTCCAA 775
Qy 721 TACTTAGTAAATAGGCTTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db 776 TACTTAGTAAATAGGCTTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAA 835
Qy 781 ATTGACGAGAGTTTAAACAGAAAGCTTGGCCAGTTGGCTCATTTGGAATGAATTCAT 840
Db 836 ATTGACGAGAGTTTAAACAGAAAGCTTGGCCAGTTGGCTCATTTGGAATGAATTCAT 895
Qy 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATCTGTGGAACTTGAATTC 900
Db 896 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATCTGTGGAACTTGAATTC 955
Qy 901 AAGGTTTTTCAGGCAGAAATCCTTTTGTATTTTATGGAACATTTCTTTTGAAGGAAA 960
Db 956 AAGGTTTTTCAGGCAGAAATCCTTTTGTATTTTATGGAACATTTCTTTTGAAGGAAA 1015
Qy 961 ACAAAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGAAGAAC 1020
Db 1016 ACAAAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGAAGAAC 1075
Qy 1021 ACGATAACGCTTCTCACTTTGATGATGAGATTTT 1053
Db 1076 ACAGATAACGCTTCTCACTTTGATGATGAGATTTT 1108
```

RESULT 3

```
AAF32439
ID AAF32439 standard; cDNA; 4955 BP.
XX
AC AAF32439;
XX
DT 18-APR-2001 (first entry)
XX
DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:3.
XX
KW Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
XX
OS Homo sapiens.
XX
PN W0200100799-A1.
XX
PD 04-JAN-2001.
XX
PF 27-JUN-2000; 2000WO-JP004189.
XX
PR 28-JUN-1999; 99JP-00181131.
PR 06-JUL-1999; 99JP-00192391.
PR 21-JAN-2000; 2000JP-00017770.
XX
(TAKE ) TAKEDA CHEM IND LTD.
(PA (NAKA/) NAKAMURA Y.
XX
PI Nakamura Y, Arakawa H, Tanaoka H;
XX
WPI; 2001-112446/12.
XX
R Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
diagnosis, treatment and prevention of cancer.
XX
Example 2; Page 87-90; 102pp; Japanese.
XX
The present invention describes a human ribonucleotide reductase
designated TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of
the DNA repair mechanism and its activity is induced by p53. It can be
used for the treatment, prevention and diagnosis of a wide range of
cancers. The present sequence represents a human ribonucleotide reductase
related sequence which is used in an example from the present invention
XX
SQ Sequence 4955 BP; 1526 A; 803 C; 999 G; 1627 T; 0 U; 0 Other;
```

```
Query Match 100.0%; Score 1053; DB 4; Length 4955;
Best Local Similarity 100.0%; Pred. No. 2.1e-290;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 245 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
Qy 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 305 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
Qy 121 CGGTTTGTCTATCTTCCAAATCCAGTACCCCTGATATTTGGAAATGTATAAACAGGCACAG 180
Db 365 CGGTTTGTCTATCTTCCAAATCCAGTACCCCTGATATTTGGAAATGTATAAACAGGCACAG 424
Qy 181 GCTTCCTTCTGGACAGCAGAAAGGTGCACTTATCAAGAGATCTCCCTCACTGGAAACAAG 240
Db 425 GCTTCCTTCTGGACAGCAGAAAGGTGCACTTATCAAGAGATCTCCCTCACTGGAAACAAG 484
Qy 241 CTTAAAGCAGATCAGAAAGTACTTCACTCTCTCACTCTTAGCCTTTTTCAGCCAGTGAT 300
Db 485 CTTAAAGCAGATCAGAAAGTACTTCACTCTCTCACTCTTAGCCTTTTTCAGCCAGTGAT 544
Qy 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAAGTTCCAGAGGCT 360
Db 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAAGTTCCAGAGGCT 604
Qy 361 CGCTGTTTCTATGCGCTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 420
```

Db 605 CGCTGTTTCTATGGCTTTCAAAATCTCATCAGAAATGTTCACTCAGAGATGTACAGTTG 664
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAATTTTATTAATGCAATGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAATTTTATTAATGCAATGAA 724
Qy 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAAA 784
Qy 541 TCTACTTTTGGGAAAGAGTGGGCTTTGCTGTAGAGAGATTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGGGCTTTGCTGTAGAGAGATTTTCTTCTCAGGA 844
Qy 601 TCTTTTGTCTATATCTGGCTTAAGAAAGAGAGTCTTATGCCAGGACTCCTTTTCC 660
Db 845 TCTTTTGTCTATATCTGGCTTAAGAAAGAGAGTCTTATGCCAGGACTCCTTTTCC 904
Qy 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGGCTGTATGTTCCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGGCTGTATGTTCCAA 964
Qy 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAAA 780
Db 965 TACTTAGTAAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAAA 1024
Qy 781 ATTGACGAGAGTTTAAACAGAAAGCTTGCAGAGTGGCCCTCATTTGGAATGAATTCATT 840
Db 1025 ATTGACGAGAGTTTAAACAGAAAGCTTGCAGAGTGGCCCTCATTTGGAATGAATTCATT 1084
Qy 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGAACTTTGGATTTCTCA 900
Db 1085 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGAACTTTGGATTTCTCA 1144
Qy 901 AAGGTTTTTTCAGGAGAGAAATCCTTTTGTATTTTATGGAAAAATTTCTTTTAGAAGGAAAA 960
Db 1145 AAGGTTTTTTCAGGAGAGAAATCCTTTTGTATTTTATGGAAAAATTTCTTTTAGAAGGAAAA 1204
Qy 961 ACAATTTCTTTGAGAAAACAGTTTTCAGATATCAGGTTTTTTCAGATTTGAGGAGAAACC 1020
Db 1205 ACAATTTCTTTGAGAAAACAGTTTTCAGATATCAGGTTTTTTCAGATTTGAGGAGAAACC 1264
Qy 1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053
Db 1265 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1297

RESULT 4

ID ADR24210
ADR24210 standard; DNA; 4955 BP.

XX
XX AC
XX AC

DT 21-OCT-2004 (first entry)

XX
DE Breast cancer prognosis marker #71.

XX
KW db; breast cancer; prognosis; gene expression; diagnosis.

XX
OS Homo sapiens.

XX
FN W02004065545-A2.

XX
PD 05-AUG-2004.

XX
PF 15-JAN-2004; 2004WO-US001100.

XX
PR 15-JAN-2003; 2003US-00342887.

XX
PA (ROSE-) ROSETTA INPHARMATICS LLC.
PA (NECA-) NETHERLANDS CANCER INST.

PI Van't Veer LJ, He Y;
XX WPI; 2004-593473/57.
XX
XX Classifying a breast cancer patient according to prognosis comprises
PT determining the similarity between the level of expression of each of
PR five genes in a cell sample taken from patient, to control levels.
XX
XX Disclosure; SEQ ID NO 71; 226pp; English.
XX
XX The invention relates to a method of classifying a breast cancer patient
CC according to prognosis by determining the similarity between the level of
CC expression of each of five genes for which markers are listed in the
CC specification, in a cell sample taken from the breast cancer patient, to
CC control levels of expression for each respective five genes to obtain a
CC patient similarity value. The methods are useful for classifying a breast
CC cancer patient according to prognosis. Kits and computer program products
CC are useful for data analysis using the diagnostic, prognostic and
CC statistical methods of the invention. This sequence corresponds to a
XX marker used in the method of the invention.
XX
SQ Sequence 4955 BP; 1526 A; 803 C; 999 G; 1627 T; 0 U; 0 Other;

Query Match 100.0%; Score 1053; DB 13; Length 4955;

Best Local Similarity 100.0%; Pred. No. 2.1e-290;

Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 245 ATGGCGGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
Qy 61 GACACCAACCAAGTGAATTAAGTCAAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 120
Db 305 GACACCAACCAAGTGAATTAAGTCAAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 364
Qy 121 CGGTTGTCATCTTTTCCAAATCCAGTACCTGATATTTGAAAAATGTATAAACAGGCACAG 180
Db 365 CGGTTGTCATCTTTTCCAAATCCAGTACCTGATATTTGAAAAATGTATAAACAGGCACAG 424
Qy 181 GCTTCTTCTGGACAGCAGAGAGTGCATTAATCAAGAGATCTCCCTCACTGGAAACAG 240
Db 425 GCTTCTTCTGGACAGCAGAGAGTGCATTAATCAAGAGATCTCCCTCACTGGAAACAG 484
Qy 241 CTTAAAGCAGATGAGAGTACTTCACTCTCACATCTTAGCCTTTTTCGAGCCAGTGTAT 300
Db 485 CTTAAAGCAGATGAGAGTACTTCACTCTCACATCTTAGCCTTTTTCGAGCCAGTGTAT 544
Qy 301 GGAATTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGCAGGTTCCAGAGCT 360
Db 545 GGAATTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGCAGGTTCCAGAGCT 604
Qy 361 CGCTGTTTCTATGGCTTTTCAAAATTTCTATCAGAGAAATGTTTCACTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTTCTATGGCTTTTCAAAATTTCTATCAGAGAAATGTTTCACTCAGAGATGTACAGTTTG 664
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAATTTTATTAATGCAATGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAATTTTATTAATGCAATGAA 724
Qy 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAAA 784
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCCCTTTGCTGTAGAGAGATTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGGTGGCCCTTTGCTGTAGAGAGATTTTCTTCTCAGGA 844
Qy 601 TCTTTTGTCTATATTTCTGGCTTAAGAAAGAGAGTCTTATGCCAGGACTCCTTTTCC 660
Db 845 TCTTTTGTCTATATTTCTGGCTTAAGAAAGAGAGTCTTATGCCAGGACTCCTTTTCC 904
Qy 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGGCTGTATGTTCCAA 720

Db 905 AATGAACATCATCAGCAGAGATGAAGGACCTTCACTGTGACCTTGTGCTGCCTGATGTTCCAA 964
Qy 721 TACTTAGTAAATAAGCCTTCAGAGAAGAGGCTCAGGAGATCATTTGTGATGCTGTCAAA 780
Db 965 TACTTAGTAAATAAGCCTTCAGAGAAGAGGCTCAGGAGATCATTTGTGATGCTGTCAAA 1024
Qy 781 ATTGAGCAGAGATTTTAAACAGAGGCTTGCAGGCTCCTCAATTTGGAATGAATTCATT 840
Db 1025 ATTGAGCAGAGATTTTAAACAGAGGCTTGCAGGCTCCTCAATTTGGAATGAATTCATT 1084
Qy 841 TTGATGAACAGTACATGAGTTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 900
Db 1085 TTGATGAACAGTACATGAGTTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 1144
Qy 901 AAGGTTTTTTCAGGAGAGAAATCCTTTTGTATTTTATGGAACATTTCTTTAGAGGAAAA 960
Db 1145 AAGGTTTTTTCAGGAGAGAAATCCTTTTGTATTTTATGGAACATTTCTTTAGAGGAAAA 1204
Qy 961 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1020
Db 1205 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1264
Qy 1021 ACAGATAACGCTCTTCACTTGGATGCGAGATTTT 1053
Db 1265 ACAGATAACGCTCTTCACTTGGATGCGAGATTTT 1297

RESULT 5

AAF32447
ID AAF32447 standard; cDNA; 1053 BP.
AC AAF32447;
DT 18-APR-2001 (first entry)
XX Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:12.
DE Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
XX Homo sapiens.
XX WO200100799-A1.
XX 04-JAN-2001.
XX 27-JUN-2000; 2000MO-JP004189.
XX 28-JUN-1999; 99JP-00181131.
PR 06-JUL-1999; 99JP-00192391.
PR 21-JAN-2000; 2000JP-0001770.
XX (TAKE) TAKEDA CHEM IND LTD.
PA (NAKA/) NAKAMURA Y.
PA Nakamura Y, Arakawa H, Tanaka H;
XX WPI; 2001-112446/12.
DR Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
PT diagnosis, treatment and prevention of cancer.
PT Claim 6; Page 96-97; 102pp; Japanese.
PS The present invention describes a human ribonucleotide reductase
XX designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of
CC the DNA repair mechanism and its activity is induced by p53. It can be
CC used for the treatment, prevention and diagnosis of a wide range of
CC cancers. The present sequence represents a human ribonucleotide reductase
CC related sequence which is given in the exemplification of the present
XX invention
XX Sequence 1053 BP; 314 A; 191 C; 246 G; 302 T; 0 U; 0 Other;

Query Match 99.8%; Score 1051.4; DB 4; Length 1053;
Best Local Similarity 99.9%; Pred. No. 2.9e-290;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGGGCGACCGGAAAGCCGGAAGCGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Db 1 ATGGGCGACCGGAAAGCCGGAAGCGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Qy 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Qy 121 CGGTTTGTCTCTTTTCCAATCCAGTACCTCTGATATTTGGAATATGTAATAACAGGCACAG 180
Db 121 CGGTTTGTCTCTTTTCCAATCCAGTACCTCTGATATTTGGAATATGTAATAACAGGCACAG 180
Qy 181 GCTTCTCTTTCGACAGCAGAGAGGTGACCTTATCAAAAGGATCTCCCTCACTGGAAACAAG 240
Db 181 GCTTCTCTTTCGACAGCAGAGAGGTGACCTTATCAAAAGGATCTCCCTCACTGGAAACAAG 240
Qy 241 CTTAAAGCAGATGAGAGTACTTCACTCTCACAATTTAGCCCTTTTTCAGCCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAGTACTTCACTCTCACAATTTAGCCCTTTTTCAGCCAGTGAT 300
Qy 301 GGAATTTGTAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGTTCCAGAGGCT 360
Db 301 GGAATTTGTAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGTTCCAGAGGCT 360
Qy 361 CGCTGTTTCTATGCTTTTCAAATTTCTATCAGAAATGTTTCACTCAGAGATGTACAGTTG 420
Db 361 CGCTGTTTCTATGCTTTTCAAATTTCTATCAGAAATGTTTCACTCAGAGATGTACAGTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATCAATGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATCAATGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAAAGAGATTTGGGCTTTCCGATGATGATGATGATGATGAT 540
Db 481 ACCATGCCCTATGTTAAGAAAAAGAGATTTGGGCTTTCCGATGATGATGATGATGATGAT 540
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTCTGCTGTAGAGAGTGTCTTCTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTCTGCTGTAGAGAGTGTCTTCTCTCAGGA 600
Qy 601 TCTTTTGTCTATATTTCTGCTAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Db 601 TCTTTTGTCTATATTTCTGCTAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Qy 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGTGATGATGATGAT 720
Db 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGTGATGATGATGAT 720
Qy 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTGATGCTGTCAAA 780
Db 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTGATGCTGTCAAA 780
Qy 781 ATTGAGCAGAGATTTTAAACAGAGGCTTGCAGGCTCCTCAATTTGGAATGAATTCATT 840
Db 781 ATTGAGCAGAGATTTTAAACAGAGGCTTGCAGGCTCCTCAATTTGGAATGAATTCATT 840
Qy 841 TTGATGAACAGTACATGAGTTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 900
Db 841 TTGATGAACAGTACATGAGTTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 900
Qy 901 AAGGTTTTTTCAGGAGAGAAATCCTTTTGTATTTTATGGAACATTTCTTTAGAGGAAAA 960
Db 901 AAGGTTTTTTCAGGAGAGAAATCCTTTTGTATTTTATGGAACATTTCTTTAGAGGAAAA 960
Qy 961 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1020
Db 961 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1020
Qy 1021 ACAGATAACGCTCTTCACTTGGATGCGAGATTTT 1053

```

Db      1021 ACAGATAACGCTCTTCACCTTGGATGACAGATTTT 1053
|||||
RESULT 6
AAF32440
ID      AAF32440 standard; cDNA; 1081 BP.
XX
AC      AAF32440;
XX
DT      18-APR-2001 (first entry)
XX
DE      Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:4.
XX
KW      Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
XX
OS      Homo sapiens.
XX
PN      W0200100799-A1.
XX
PD      04-JAN-2001.
XX
PF      27-JUN-2000; 2000WO-JP004189.
XX
PR      28-JUN-1999; 99JP-00181131.
XX
PR      06-JUL-1999; 99JP-00192391.
XX
PR      21-JAN-2000; 2000JP-00017770.
XX
(PAKE ) TAKEDA CHEM IND LTD.
PA      (NAKA/) NAKAMURA Y.
XX
PI      Nakamura Y, Arakawa H, Tanaka H;
XX
XX      WPI; 2001-112446/12.
XX
DR      Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
PT      diagnosis, treatment and prevention of cancer.
XX
XX      Example 2; Page 91; 102pp; Japanese.
XX
The present invention describes a human ribonucleotide reductase
CC      designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of
CC      the DNA repair mechanism and its activity is induced by p53. It can be
CC      used for the treatment, prevention and diagnosis of a wide range of
CC      cancers. The present sequence represents a human ribonucleotide reductase
CC      related sequence which is used in an example from the present invention
XX
XX      Sequence 1081 BP; 322 A; 199 C; 252 G; 308 T; 0 U; 0 Other;
SQ
Query Match          99.8%; Score 1051.4; DB 4; Length 1081;
Best Local Similarity 99.9%; Pred. No. 2.9e-290;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1  ATGGGCGACCCGGAAAGCGCGGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db      |||||||
QY      20  ATGGGCGACCCGGAAAGCGCGGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 79
Db      |||||||
QY      61  GACACCAACGAAAGTGAATTAAGTCAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Db      |||||||
QY      80  GACACCAACGAAAGTGAATTAAGTCAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 139
Db      |||||||
QY      121 CGGTTGTCTATCTTTCCATCCAGTACCTCGATATTTGGAAATGTATTAACAGGACAG 180
Db      |||||||
QY      140 CGGTTTGTCTATCTTTCCATCCAGTACCTCGATATTTGGAAATGTATTAACAGGACAG 199
Db      |||||||
QY      181 GCTTCCTCTCGGACAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGAAACAAG 240
Db      |||||||
QY      200 GCTTCCTCTCGGACAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGAAACAAG 259
Db      |||||||
QY      241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCAATCTTAGCCCTTTTTCAGGCCAGTGAT 300
Db      |||||||
QY      260 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCAATCTTAGCCCTTTTTCAGGCCAGTGAT 319

```

```

QY      301 GGAATTCGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
Db      |||||||
QY      320 GGAATTCGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 379
Db      |||||||
QY      361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db      |||||||
QY      380 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 439
Db      |||||||
QY      421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTATTAATGCAATTGAA 480
Db      |||||||
QY      440 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTATTAATGCAATTGAA 499
Db      |||||||
QY      481 ACCATGCCCTATGTTAAGAAAAAGAGATTTGGCCCTTCGATGATGATGATGATGATGATGAT 540
Db      |||||||
QY      500 ACCATGCCCTATGTTAAGAAAAAGAGATTTGGCCCTTCGATGATGATGATGATGATGATGAT 559
Db      |||||||
QY      541 TCTACTTTTGGGGAAGAGTGGTGGCTTGGCTGCTGTGTAGAAAGGAGTTTCTTCTCAGGA 600
Db      |||||||
QY      560 TCTACTTTTGGGGAAGAGTGGTGGCTTGGCTGCTGTGTAGAAAGGAGTTTCTTCTCAGGA 619
Db      |||||||
QY      601 TCTTTTGTCTATATTTCTGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTCTCC 660
Db      |||||||
QY      620 TCTTTTGTCTATATTTCTGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTCTCC 679
Db      |||||||
QY      661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGGCTGCTGATGATGATGATGAT 720
Db      |||||||
QY      680 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGGCTGCTGATGATGATGATGAT 739
Db      |||||||
QY      721 TACTTAGTAATAAGCCTTCAGAAAGAGGTCTAGGAGATCATTTGTTGATGCTGTGCAAA 780
Db      |||||||
QY      740 TACTTAGTAATAAGCCTTCAGAAAGAGGTCTAGGAGATCATTTGTTGATGCTGTGCAAA 799
Db      |||||||
QY      781 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGGCAGTTCGCTCATTTGGAAATGAATTGCATT 840
Db      |||||||
QY      800 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGGCAGTTCGCTCATTTGGAAATGAATTGCATT 859
Db      |||||||
QY      841 TTGATGAAACAGTACATTTGATGTTGATGCTGACAGATTTACTTTGTGAACTTGGATTTCTCA 900
Db      |||||||
QY      860 TTGATGAAACAGTACATTTGATGTTGATGCTGACAGATTTACTTTGTGAACTTGGATTTCTCA 919
Db      |||||||
QY      901 AAGGTTTTTCAGGCGAGAAATTCCTTTTGAATTTTATGGAATAATTTCTTTAGAGGAAAA 960
Db      |||||||
QY      920 AAGGTTTTTCAGGCGAGAAATTCCTTTTGAATTTTATGGAATAATTTCTTTAGAGGAAAA 979
Db      |||||||
QY      961 ACAATTTCTTTGAGAAACGAGTTTCAGATATCAGCGTTTTCAGATTTATGCGAGAAACC 1020
Db      |||||||
QY      980 ACAATTTCTTTGAGAAACGAGTTTCAGATATCAGCGTTTTCAGATTTATGCGAGAAACC 1039
Db      |||||||
QY      1021 ACAGATAACGCTCTTCACCTTGGATGACAGATTTT 1053
Db      |||||||
QY      1040 ACAGATAACGCTCTTCACCTTGGATGACAGATTTT 1072
Db      |||||||
RESULT 7
AAH14924
ID      AAH14924 standard; cDNA; 1601 BP.
XX
AC      AAH14924;
XX
DT      26-JUN-2001 (first entry)
XX
DE      Human cDNA sequence SEQ ID NO:12810.
XX
KW      Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS      Homo sapiens.
XX
PN      EP1074617-A2.
XX
PD      07-FEB-2001.
XX
PF      28-JUL-2000; 2000EP-00116126.
XX

```


Qy	89	ATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTTCCAAATCCAGTACC	148
Db	283	AGGATGAGCGCGTCTGCTGAGAGAAACCCCGCGCGTTTGTTCATCTTTCCCAATCGAGTACC	342
Qy	149	CTGATATTTTCGAAAAATGTATAAACAGGACACAGGCTTCCTTCGACAGCAGAGAAGAGTGC	208
Db	343	ATGATATCTGGCAGATGTATTAAGAGCAGAGGCTTCCTTTTGGACCGCCGAGAGGTGG	402
Qy	209	ACTTATCAAAAGGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTACTTTCATCT	268
Db	403	ACCTCTCCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT	462
Qy	269	CTCACATCTTAGCGCTTTTTCGACCCAGTGATGGAAATGTAAATGAAAAATTTGGTGGAGC	328
Db	463	CCCATGTTCTGGGCTTCTTTTCGACAAAGCATGGCATAGTAAATGAAAAATTTGGTGGAGC	522
Qy	329	GC'TTTAGTCAGGAGGTGCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAAATCTCA	388
Db	523	GA'TTTAGCCAAAGTTTCAGATTACAGNAGCCGCTGT'TTCTATGGCTTCCAAATTCGCCA	582
Qy	389	TCGAGAAATGTTCACTCAGAGATGTACAGTTTGTGTATAGACACTTACATCAGAGATCCCA	448
Db	583	TGGAAAAACATACATTTCTGAAATGTATAGTCTTCTTATATGACACTTACATAAAAAGATCCCA	642
Qy	449	AGAAAAAGGAATTTTATATTTAAATGCAATTCGAACCAATGCCCTATGTTTAAAGAAAAAGCAG	508
Db	643	AAGAAAGGGAATTTTCTCTTCAATGCCAATGAAACGATGCCCTGTGTGTCAGAGAGGCAG	702
Qy	509	ATTGGGCTTTGCGATGGATAGCAGATAGAAAAATCTACTTTTGGGGAAAGAGTGGTGGCCT	568
Db	703	ACTGGGCTTTGCGCTGA'TTTGGGACAAAGAGGCTACCTATGTTGAACGTTGTGTAGCCT	762
Qy	569	TTGCTGCTGTAGAAAGAGTTTTCCTTCAGAGATCTTTTGTCTGCTATATTTCTGGCTAAAGA	628
Db	763	TTGCTGCAGTCGAAGGCAATTTCTTTTCCGGTTCCTTTGCGTGCATATTTCTGGCTCAAGA	822
Qy	629	AGAGAGTCTTATGSCCAGGACTCACTTTTCCATGAACCTCATCAGCAGAGATGAAAGAC	688
Db	823	AACGAGACTGTATGCTTGGCTCACA'TTTTCTAATGAACCTTATTAGCAGAGATGAGGTT	882
Qy	689	TTCACTGTGACTTTGCTTGGCTGATGTTCCAATACTTAGTAAATTAAGCCTTCAGAAGAAA	748
Db	883	TACACTGTGATTTTGTCTTGCCTGATGTTCAACACACCTTGGTACACAAACCATCGAGGAGA	942
Qy	749	GGGTACGGAGATCA'TTGTGTGCTGTCAAAATTGACGAGGAGTTT'TTAAACAGAGCCT	808
Db	943	GAGTAAAGAGAAATAATTTATCAATGCTGTTCGGATAGAACAGGAGTTCCTCAGTGAAGCCT	1002
Qy	809	TGCCAGTTTGGCTCATTTGGAAATGAATTTGCA'TTTTGAATGAACACAGTACATTTGAGTTTGTAG	868
Db	1003	TGCCTGTGAAGCTCATTTGGGATGNA'TTGGCACTCTAATGAAGCAATACATTTGAGTTTGTGG	1062
Qy	869	CTGACAGATTACTTGTGGAAC'TTGGATTCCTCAAAAGTTTTCAGGCGAGAAAAATCCTTTTG	928
Db	1063	CAGACACATTA'TGCTGGAAC'TGGT'TTTTAGCAAGT'TTTTTCAGAGTAGAGAACCACTTG	1122
Qy	929	ATTTTATGGAAGAACATTTCTTTAGAAAGGAAAAACAAATTTCTTTTGAGAAACGAGTTTCAG	988
Db	1123	ACTTATAGGAATATTTTCACTTGGAAAGAAAGACTTAACTCTTTTGGAAAGAGAGTAGCGC	1182
Qy	989	AGTATCAGCGTTTTCAGTTTATGGCAGAAAAACACAGATAACGCTTCTCACCTTGGATGCAG	1048
Db	1183	AGTATCAGAGGATGGGAGTGATGTCAGTCCAAACAGAGAA'TTCTTTTACCTTGGATGCTG	1242
Qy	1049	ATTT 1052	
Db	1243	ACTT 1246	

RESULT 11
ADJ56536
ID ADJ5
XX

AC	ADU56536;
XX	
XX	06-MAY-2004 (first entry)
XX	
DE	Murine cDNA differentially expressed in MYCN activated cells SeqID 342.
XX	
KW	mouse; murine; differential expression; transactivator; proto-oncogene;
KW	neuroblastoma; small cell lung cancer; cycostatic; gene therapy; ss;
KW	MYCN activated cell.
XX	
OS	Mus musculus.
XX	
XX	US2003119009-A1.
PN	
XX	
PD	26-JUN-2003.
XX	
PF	25-FEB-2002; 2002US-00084817.
XX	
PR	23-FEB-2001; 2001US-0270784P.
XX	
PA	(STUA/) STUART S G.
PA	(NUCH/) NUCHTERN J G.
PA	(PLON/) PLON S E.
PA	(SHOH/) SHOHEIT J M.
XX	
XX	Stuart SG, Nuchtern JG, Plon SE, Shohet JM;
PI	
XX	WPI; 2003-635698/60.
DR	
XX	
XX	New genes regulated by MYCN activation, useful in gene therapy.
PT	particularly for treating a subject with e.g. neuroblastoma or other
PT	cancers, or for diagnosing, staging or monitoring the treatment of the
PT	cancer.
XX	
XX	Claim 1; SEQ ID NO 342; 27pp; English.
XX	
XX	This invention relates to novel isolated cDNAs that are differentially
CC	expressed in MYCN activated cells. Specifically, it refers to
CC	polynucleotide sequences that exhibit differential expression patterns in
CC	cells activated by the transactivator MYCN, where MYCN is a proto-
CC	oncogene that is amplified in neuroblastoma cells and is common in small
CC	cell lung cancers. The present invention describes these cDNA molecules
CC	as useful for in hybridisation assays to detect expression of nucleic
CC	acids (or complementary nucleic acids) in a present in a given sample, as
CC	well as for screening assays by identifying molecules or compounds that
CC	specifically bind the cDNA as a ligand and modulate function or activity.
CC	Accordingly, these compositions exhibit cytostatic activity and can also
CC	be used for gene therapy purposes. This polynucleotide sequence is a cDNA
CC	that is differentially expressed in MYCN activated cells, given in an
CC	exemplification of the invention. NOTE: This sequence does not appear in
CC	the printed specification but has been obtained in electronic format from
CC	the US Patent Office at
CC	ftp.segdata.uspto.gov/sequence.html?DocID=20030119009.
XX	
XX	Sequence 2216 BP; 593 A; 476 C; 524 G; 623 T; 0 U; 0 Other;

!

Db 1966 ACCTCTCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 1907
Qy 269 CTACATCTTTAGCCTTTTTCGAGCCAGTGAATTTGTAATGAAATTTGTTGGAGC 328
Db 1906 CCATGTTCTGGCTTTCTTTCGAGCAAGCGATGGCATAGTAATGAAATCTTGGTGAGC 1847
Qy 329 GCTTTAGTCAGGAGTCCAGGTTCCAGAGCGTGGCTTTCTATGGCTTTCAAAATTTCTCA 388
Db 1846 GATTTAGCCAGAGTTTCAGATTACAGAGCCCGCTGTTTCTATGGCTTCCAAATTTGCCA 1787
Qy 389 TCAGAAATGTTCACTACAGAGATGACAGTTTGTGTATAGACATTTACATCAGAGATCCCA 448
Db 1786 TGGAAACATACATCTTGAATATGATAGTCTTCTTATGACACTTACATAAAGATCCCA 1727
Qy 449 AGAAAGGGAAATTTTATTTAAATGCAATTCGAACCAATGCTATGTTTGAAGAAAGACAG 508
Db 1726 AAGAAAGGGAAATTTCTCTTCAATGCGCAATGAAACGATGCTTGTGTCAAGAAGACAG 1667
Qy 509 ATTGGGCTTGGCATGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCCT 568
Db 1666 ACTGGGCTTGGCTGATTTGGGCAAGAGGCTACTATGTTGAACGTGTTGTAGCCT 1607
Qy 569 TTGCTGCTGAGAGGAGTTTCTTCTCAGAGATCTTTTGTGCTATATTTCTGCTAAAGA 628
Db 1606 TTGCTGAGTGAAGGCAATTTCTTTCCGGTCTTCTTGGCTCGATATTTCTGGCTCAAGA 1547
Qy 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCATGAATCACTCAGCAGAGATGAAGGAC 688
Db 1546 AACGAGGACTGATGCGCTGCGCTCAATTTTCTAATGAATTAATAGCAGAGATGAGGTT 1487
Qy 689 TTCACTGCTGACTTTGCTGCTGATGTTTCAATCTTCTAGTAATTAAGCCCTTCAGAGAAA 748
Db 1486 TACACTGTGATTTTGTCTGCTGATGTTTCAACCTGTTACAAACCAATCGGAGAGA 1427
Qy 749 GGGTCAGGAGATCAATTTGTTGATGCTCTCAAAATTTGAGCAGGAGTTTTCACAGAACCT 808
Db 1426 GAGTAAGAGAAATTAATTAATCAATGCTGCTCGGATAGAACAGGAGTTCTCTCAGGCGCT 1367
Qy 809 TGCAGTTGGCCTCAATTTGGAATGATGATTTGATGAAACAGATACATTTGATTTGTTAG 868
Db 1366 TGCTGTGAAGCTCAATTTGGGATGAATTTGCACTCTAATGAAGCAATACATTTGATTTG 1307
Qy 869 CTGACAGATTTCTGCTGAACTCTGATTTCTCAAGGTTTTCAGGCAAGAAATCCTTTTG 928
Db 1306 CAGACAGATTTGCTGAACTGCTGTTTGGAGGTTTTCAGAGTAGAAGCAACCAATTTG 1247
Qy 929 ATTTTATGAAAAATTTCTTTAGAGGAAAAACAAATTTCTTTGAGAAACGAGTTTTCAG 988
Db 1246 ACTTTATGAGAAATTTCTCTGGAAGGAAAGACTAATCTTTTGAAGAGAGTAGGCG 1187
Qy 989 AGTATCAGGTTTTCAGTTTATGCGAGAAACCAAGATACGTTCTTCACTTGGATGAG 1048
Db 1186 AGTATCAGAGATGGGAGTGTGATGTCAGATGCAACAGAGAAATTTCTTTACCTTGGATGCTG 1127
Qy 1049 ATTT 1052
Db 1126 ACTT 1123

RESULT 13
ABL65414
ID ABL65414 standard; DNA; 2500 BP.
XX AC ABL65414;
XX AC
XX AC
DT 15-MAY-2002 (first entry)
XX Lung cancer related gene sequence SEQ ID NO:3751.
DE Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW Gene; ds.

XX Homo sapiens.
OS WO200194629-A2.
PN 13-DEC-2001.
PD 30-MAY-2001; 2001WO-US010838.
XX 05-JUN-2000; 2000US-0209473P.
XX 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 01-NOV-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
PA (AVAL-) AVALON PHARM.
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.
XX Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX Claim 1; SEQ ID NO 3751; 44pp; English.
XX The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an

CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX

SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;

Query Match 54.9%; Score 578.4; DB 6; Length 2500;
Best Local Similarity 75.0%; Pred. No. 1.2e-154;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCTCTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTTCCATTCAGTACC 148
DB 397 AGGATGAGCGCTGCTGAGAGAAAACCCCGCGCTTTGTCTATCTTTCCCATCGAGTACC 456
QY 149 CTGATATTTGGAAAATGTATAACAGGCACAGSCTTCTCTCGCAGCAGAGAGGTGG 208
DB 457 ATGATATCTGGCAGATGATAGAAGGCAGAGGCTTCTTTTGGACCGCGAGGAGGTGG 516
QY 209 ACTTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTCATCT 268
DB 517 ACCTCTCAAGGACATTCAGCACTGGATCCCTGAAACCCGAGGAGATATTTATAT 576
QY 269 CTCACATCTTAGCTTTTGGCAGCAGTGATGGAATTTGAAATGAAATTTGGTGGAGC 328
DB 577 CCCATGTTCTGGCTTCTTTTGCAGCAAGCGATGGCAGATAGTAAATGAAACTTTGGTGGAGC 636
QY 329 GCTTTAGTCAGAGTGAGGTTCCAGAGGCTGCGTGTCTTCTATGCGTTCGAAATCTCA 388
DB 637 GATTTAGCCAAAGGTCAGATTACAGAAAGCGCGTGTCTTCTATGCGCTTCCAAATTTGCCA 696
QY 389 TCGAGAAATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCA 448
DB 697 TCGAAAACATACATCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 756
QY 449 AGAAAAGGAAATTTTATTTAAATGCAATTTGAACCATGCGCTATGTTTAAAGAAAAGCAG 508
DB 757 AAGAAAAGGAAATTTCTCTTCAATGCTTGAACCATGCGCTTGTGTCAGAGAGAGGCGAG 816
QY 509 ATTGGCCCTTGGATGATGATGAGATGAAATCTACTTTTGGGGAAGAGTGGTGGCT 568
DB 817 ACTGGCCCTTGGCTGGATTTGGGGAACAAAGAGGCTACCTATGCTGAAACGTTGTAGCCT 876
QY 569 TTGCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTCTGCTATATTTCTGGCTAAAGA 628
DB 877 TTGCTGCTGATGAGGAGTTTCTTCTTCCGGTCTTTTGGCTGATATTTCTGGCTCAAGA 936
QY 629 AGAGAGGCTTTATGCGCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGAC 688
DB 937 AACGAGGACTGATGCTGCGCTCACATTTCTTAATGAACCTTATAGCAGAGATGAGGTT 996
QY 689 TTCACTGTGACTTTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 748
DB 997 TACACTGTGATTTTGTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1056
QY 749 GGGTCAGGAGATCATCTGATGCTGCAAAATTTGAGCAGGAGTTTAAACAGAGCT 808
DB 1057 GAGTAAAGAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1116
QY 809 TGCCAGTTGGCTCATTTGGAATGAAATTTGATGAAACAGTACATGATGATGATGATGAT 868
DB 1117 TGCCAGTTGGCTCATTTGGAATGAAATTTGATGAAACAGTACATGATGATGATGATGAT 1176
QY 869 CTGACAGATTAATCTGAGGACTTTGATGATGATGATGATGATGATGATGATGATGATGAT 928
DB 1177 CAGACAGACTTATGCTGGAACCTGGGTTTTAGCAAGGTTTTTACAGATGAGAGACCATTTG 1236
QY 929 ATTTTATGAAAACATTTCTTTAGAGGAAACAAATTTCTTTTGGAGAACAGGTTTTCAG 988

DB 1237 ACTTTATGAGAAATTTTCACTGGAAGAAAGACTAACTTCTTTTGAAGAGAGTAGCGG 1296
QY 989 AGTATCAGCGTTTTCAGAGTATGGCAGAAACCCAGATAAGCTTTCACCTTGGATGAG 1048
DB 1297 AGTATCAGAGGATGGGAGTGTCAAGTCCAAACAGAGAAATTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
DB 1357 ACTT 1360

RESULT 14
ABL66517
ID ABL66517 standard; DNA; 2500 BP.
XX ABL66517;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:4854.
XX
KW Human; cancer; colon; breast; kidney; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.

PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 4854; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL6164
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms
CC tumour
XX
SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;

Query Match 54.9%; Score 578.4; DB 6; Length 2500;
Best Local Similarity 75.0%; Pred. No. 1.2e-154;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTGTGTCATCTTTCCAAATCCAGTACC 148
Db AGGATGAGCGGCTGTGAGAGAAACCCCGCGGCTGTGTCATCTTTCCCATCGAGTACC 456
QY 149 CTGATATTTGGAAATGTATAAACAGGCAAGGCTTCCTTTGGACAGAGAGAGGTGG 208
Db ATGATATCTGGCAGATGTATAAGAGGCGAGAGGCTTCCTTTGGACCGCGGAGGAGGTG 516
QY 209 ACTTATCAAGGATCTCCCTCACTGGAACAAGCTTTAAGCAGATGAGAAGTACTTCACT 268
Db ACCCTTCCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
QY 269 CTCACATCTTACGCTTTTTCAGCCAGGTGAGTAATCTTAATGAAATTTGTGGAGC 328
Db CCCATGTTCTGGCTTCTTTTGACGAGCGATGAGTAATGAGTAATGAAATTTGTGGAGC 636
QY 329 GCTTTAGTCAGGAGGTGCGGTTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTTCTCA 388
Db GATTTAGCAAGAAGTTCAGATTACAGAGCGCGCTGTTTCTATGGCTTTCAAATTTGCA 696
QY 389 TCGAAGATGTTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTATCATCAGAGATCCCA 448
Db TCGAAGAACATACATCTCGAAATGTATGTCTTTTATTCACACTTACATATAAAGATCCCA 756
QY 449 AGAAGAGGAAATTTTATTTATGCAATTTGAACCAATGCCCTATCTTAAAGAAAAGCAG 508
Db AAGAAAGGAAATTTCTTTCAATGCAATTTGAACCGATGCTTTGTGTCAAGAAAGAGGCGAG 816
QY 509 ATTGGGCTTGGCATGGATGAGCAGATAGAAAATCTACTTTTGGGGAAGAGGTGGTGGCT 568
Db ACTGGGCTTGGCATGGATGAGCAGATAGAAAATCTACTTTTGGGGAAGAGGTGGTGGCT 876

QY 569 TTGCTGCTCTAGAGGAGGTTTTTCTCTCAGGATCTTTTCTGCTGCTATATTTCTGGCTAAGA 628
Db TTGCTGCTGCTAGAGGAGGTTTTTCTCTCAGGATCTTTTCTGCTGCTATATTTCTGGCTAAGA 936
QY 629 AGAGAGGCTTTATGCGAGGAGCTCACATTTTCCAAATGAACTCATCAGCAGAGATGAAGAC 688
Db AACGAGGAGCTGATGCTGCGCTCACATTTTCTAAATGAACCTTATTAGCAGAGATGAGGCTT 996
QY 689 TTCACTGTGACATTTGCTTGCCTGATGTTCCAAATGAACTCATCAGCAGAGCTTTCAGAGAAA 748
Db TACACTGTGATTTGCTTGCCTGATGTTCCAAACACCTGGTACAAACCAATCGAGAGAGA 1056
QY 749 GGGTCAGGAGAGATCATTTGTTGATGCTGTCAAAATTCAGCAGGAGGTTTTTAAACAGAACCT 808
Db GAGTAAGAGAAATATATCAATGCTGTTCCGATAGACAGGAGTTCTCTCAGTGGGCTT 1116
QY 809 TCCAGTGTGCTCATTTGGAATGAAATTTGATGAAACAGTACATTTGAGTTTCTAG 868
Db TGCCTGTGAAGCTCATTTGGGATGAAATTTGCACTCTAATGAAGCAATACATTTGAGTTTCTGG 1176
QY 869 CTGACAGATTTACTTGTGGAACCTTGGATTTCTCAAGGTTTTTTCAGGAGAAATCCCTTTG 928
Db CAGACAGACTTATGCTGGAACCTGGGTTTTAGCAAGGTTTTTCAGAGTAGAGAACCCATTTG 1236
QY 929 ATTTTATGAAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTTCAG 988
Db ACTTTATGAGAAATATTTTCACTGGAAGGAAAGACTAACTTTCTTTGAGAGAGTAGGCG 1296
QY 989 AGTATCAGGTTTTTTCAGTTATGGCAGAAACCAAGATACAGTCTTTCACCTTGGATCAG 1048
Db AGTATCAGAGGATGGGAGTGATGTCAGATCCCAACAGAGAAATTTCTTTTACCTTGGATCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 15
ABL65859
ID ABL65859 standard; DNA; 2500 BP.
XX
AC ABL65859;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:4196.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilms's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
XX WO200194629-A2.
XX
PD 13-DEC-2001.
XX
XX 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.

PR 25-SEP-2000; 2000US-0235082P.
 PR 25-SEP-2000; 2000US-0235134P.
 PR 25-SEP-2000; 2000US-0235280P.
 PR 26-SEP-2000; 2000US-0235637P.
 PR 26-SEP-2000; 2000US-0235638P.
 PR 27-SEP-2000; 2000US-0235711P.
 PR 27-SEP-2000; 2000US-0235720P.
 PR 27-SEP-2000; 2000US-0235840P.
 PR 27-SEP-2000; 2000US-0235863P.
 PR 28-SEP-2000; 2000US-0236028P.
 PR 28-SEP-2000; 2000US-0236032P.
 PR 28-SEP-2000; 2000US-0236033P.
 PR 28-SEP-2000; 2000US-0236034P.
 PR 28-SEP-2000; 2000US-0236109P.
 PR 28-SEP-2000; 2000US-0236111P.
 PR 29-SEP-2000; 2000US-0236842P.
 PR 29-SEP-2000; 2000US-0236891P.
 PR 02-OCT-2000; 2000US-0237172P.
 PR 02-OCT-2000; 2000US-0237173P.
 PR 02-OCT-2000; 2000US-0237278P.
 PR 02-OCT-2000; 2000US-0237294P.
 PR 02-OCT-2000; 2000US-0237295P.
 PR 02-OCT-2000; 2000US-0237316P.
 PR 03-OCT-2000; 2000US-0237425P.
 PR 03-OCT-2000; 2000US-0237598P.
 PR 03-OCT-2000; 2000US-0237604P.
 PR 03-OCT-2000; 2000US-0237606P.
 PR 03-OCT-2000; 2000US-0237608P.
 PR 01-NOV-2000; 2000US-0244867P.
 PR 01-NOV-2000; 2000US-0245084P.
 XX (AVAL-) AVALON PHARM.
 XX
 XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX
 XX WPI; 2002-188264/24.
 XX
 PT Screening for anti-neoplastic agent involves exposing cells to a chemical
 PT agent to be tested for anti-neoplastic activity, and determining a change
 PT in expression of a gene of a signature gene set.
 XX
 PS Claim 1; SEQ ID NO 4196; 44pp; English.
 XX
 CC The present invention describes a method (M1) for screening for an anti-
 CC neoplastic agent. The method involves exposing cells to a chemical agent
 CC to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (II) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening an
 CC anti-neoplastic agent, and can be used for producing a product which is
 CC the data collected with respect to the anti-neoplastic agent as a result
 CC of M1, and the data is sufficient to convey the chemical structure and/or
 CC properties of the agent. M1 can be used in the treatment of cancer: such
 CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
 CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
 CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
 CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms
 CC tumour
 XX
 SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;
 Query Match 54.9%; Score 578.4; DB 6; Length 2500;
 Best Local Similarity 75.0%; Pred. No. 1.2e-154;
 Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;
 QY 89 ATGAGAGCCACTCTCTAAGAAAGAGTTCTCCCGGTTTGTTCATCTTCCAAATCCAGTACC 148
 DB |||||
 DB 397 AGGATGAGCGCTGCTGAGAGAAAACCCCGCGCTTTGTTCATCTTCCCATCGAGTACC 456
 QY 149 CTGATATTTGGAAATGTATTAACAGGCACAGGCTTCCTCTTGGACAGAGAGGTGCG 208

Db ||||| ATGATATCTGGCAGATGTATTAAGAGCAGAGGCTTCCTTTTGGACCCGCGAGGTTG 516
 QY ||||| ACTTATCAAGAGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAGTACTTCTCT 268
 Db ||||| ACCTCTCCAAGGACATTCAGACACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTAT 576
 QY ||||| CTCACATCTTAGCCTTTTGGAGCCAGTGTGGAATGTGAAATGAAATTTGGTGGAGC 328
 Db ||||| CCATGTTCTGGCTTCTTTGACGCAAGCGATGGCATAGTAATGAAATCTTGGTGGAGC 636
 QY ||||| GCTTTAGTCAGGAGGTGCAAGTTCAGAGGCTCGCTGTTCTTATGGCTTTCAAAATTTCTCA 388
 Db ||||| GATTTAGCCAAGAAGTTTCAGATTACAGAGCCGCTGTTCTTATGGCTTCCAAATTTGCCA 696
 QY ||||| TCGAGAAATGTTCTACTCAGAGATGTACAGTTTGTGTGATAGACACTTACATCAGAGATCCCA 448
 Db ||||| TCGAAACACATACATCTGAAATGTATAGTCTTCTTATTCACACTTACATAAAGATCCCA 756
 QY ||||| AGAAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCCCTATGTTAAGAAAAGAGCAG 508
 Db ||||| AGAAAAGGGAATTTCTTCAATGCAATTTGAAACCATGCCCTTGTGTCAAGAAAGGAGCAG 816
 QY ||||| ATTGGGCTTGGATGAGATAGAGATAGAAAATCTACTTTTGGGAAAGAGTGGTGGCT 568
 Db ||||| ACTGGGCTTGGCTGATTTGGGACAAAGAGGCTACCTATGTTGAACGTGTTGTAGCCT 876
 QY ||||| TTGCTCTGTAGAGGAGTTTCTTCTCAGGATCTTTTGTGCTGTATATCTCGGTAAGA 628
 Db ||||| TTGCTGAGTGGAGGCAATTTCTTTTCGGTCTTTTGGTGGATATCTTGGCTCAAGA 936
 QY ||||| AGAGAGTCTTATGCCAGGACTCACATTTTCCAAATGCAATCTATCAGCAGAGATGAGGAC 688
 Db ||||| AACGAGACTGATGCCCTGCCCTCACATTTTCTAATCAACTTATTAGCAGAGATGAGGTT 996
 QY ||||| TTCACTGTGACTTTGCTTGCCTGATGTTCCAAATCTTAGTAAATAGCCTTCAGAAGAAA 748
 Db ||||| TACACTGTGATTTGCTTGCCTGATGTTCAACACTGTTCAACAACCAATCGGAGAGA 1056
 QY ||||| GGGTCAGGAGATCAATTTGTGATGCTGTCAAAATTCAGCAGGAGTGTCTTAAACAGAGCCT 808
 Db ||||| GAGTAAGAGAAATTAATATCAATGCTGTTTCGGATAGAACAGGAGTTCCTCACTGAGCCT 1116
 QY ||||| TGCCAGTTGGCCTCATTTGGAATGAATGCAATTTTGATGAAAACAGTACATTTAGTTGTAG 868
 Db ||||| TGCCTGTGAAGCTCATTTGGGATGAATTCGACTCTAATGAAGCAATACATTTGAGTTGTGG 1176
 QY ||||| CTGACAGATTACTTGTGGAACCTTGGATCTTCAAGGTTTTCAGGCGAGAAAATCCTTTTG 928
 Db ||||| CAGACAGATTATGCTGGAACTGGGTTTTCAGCAAGGTTTTCAGAGTAGAAGCAATTTG 1236
 QY ||||| ATTTATGAAAACATTTCTTTAGAGGAAAACAAATTTCTTTGAGAAAACGAGTTTTCAG 988
 Db ||||| ACTTTATGAGATATTTCACTGGAAGGAAAGACTTAACCTTCTTTGAGAGAGATAGGCG 1296
 QY ||||| AGTATCAGGCTTTTGCAGTTATGGCAGAAAACCAAGATAACGTCCTTCACTTGGATGAG 1048
 Db ||||| AGTATCAGAGGATGGGAGTGTCAAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTG 1356
 QY ||||| ATTT 1052
 Db ||||| ACTT 1360

Search completed: October 30, 2005, 00:02:44
 Job time : 447.5 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	1053	100.0	3292	3	CR617553	CR617553 full-leng
2	1053	100.0	4748	3	HSN802033	AL137348 Homo sapi
3	1041	98.9	4650	3	BC042948	BC042948 Homo sapi
4	1010.6	96.0	3501	3	CR627376	CR627376 Homo sapi
5	945	89.7	1056	9	CR619873	CR619873 Homo sapi
6	929.6	88.3	1083	1	AL547501	AL547501 AL547501
7	829.8	78.8	1008	1	AL549857	AL549857 Pan. trogl
8	793.2	75.3	879	5	B0441857	B0441857 AGENCOURT
9	777	73.8	1043	9	CR619875	B0441857 AGENCOURT
10	731.2	69.4	892	7	CN163314	CN163314 Mus muscu
11	728.2	69.2	914	7	CF995079	CF995079 AGENCOURT
12	710.6	67.5	870	7	C0648380	C0648380 ILLUMIGEN
13	682.8	64.8	804	6	C0656661	C0656661 AGENCOURT
14	655.4	62.2	1061	4	BM801298	BM801298 AGENCOURT
15	654.8	62.2	896	5	BUI70979	BUI70979 AGENCOURT
16	616.4	58.5	685	5	BX951374	BX951374 DKFzp781E
17	578.4	54.9	1571	3	CR603461	CR603461 full-leng
18	578.4	54.9	1573	3	CR625489	CR625489 full-leng
19	578.4	54.9	1582	3	CR608076	CR608076 full-leng
20	578.4	54.9	1588	3	CR602054	CR602054 full-leng
21	578.4	54.9	1592	3	CR621427	CR621427 full-leng
22	578.4	54.9	1600	3	CR604378	CR604378 full-leng
23	578.4	54.9	1605	3	CR590959	CR590959 full-leng
24	578.4	54.9	1612	3	CR614990	CR614990 full-leng

```
QY 61 GACACCAACGAAGTGAATTAAGTCAAAATGAAGACCACTCTTAAGAAAGAGTTCTTCGC 120
DB 61 |
QY 87 GACACCAACGAAGTGAATTAAGTCAAAATGAAGACCACTCTTAAGAAAGAGTTCTTCGC 146
DB 87 |
QY 121 CGGTTTGTGATCTTTTCCAAATCCAGTACCTGTATTTTGGAAAATGATTAACACAGGCACAG 180
DB 121 |
QY 147 CGGTTTGTGATCTTTTCCAAATCCAGTACCTGTATTTTGGAAAATGATTAACACAGGCACAG 206
DB 147 |
QY 181 GCTTCCTCTTGACACGACGAAGAGTGCAGTATTAAGAGATCTCCCTCACTGGACACAG 240
DB 181 |
QY 207 GCTTCCTCTTGACACGACGAAGAGTGCAGTATTAAGAGATCTCCCTCACTGGACACAG 266
DB 207 |
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTTAGCCCTTTTTCAGCCAGTGTAT 300
DB 241 |
QY 267 CTTAAAGCAGATGAGAAGTACTTCTCATCTTAGCCCTTTTTCAGCCAGTGTAT 326
DB 267 |
QY 301 GGAATTTGTAATGAAATTTGTGGAGCGCTTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
DB 301 |
QY 327 GGAATTTGTAATGAAATTTGTGGAGCGCTTTTAGTCAGAGGTGAGGTTCCAGAGGCT 386
DB 327 |
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
DB 361 |
QY 387 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 446
DB 387 |
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTTGAA 480
DB 421 |
QY 447 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTTGAA 506
DB 447 |
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAA 540
DB 481 |
QY 507 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAA 566
DB 507 |
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGAGTCTTTCTTCTCAGGA 600
DB 541 |
QY 567 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGAGTCTTTCTTCTCAGGA 626
DB 567 |
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAAAGAGAGTCTTATGCGAGACTCACTTTTTC 660
DB 601 |
QY 627 TCTTTTGTCTATATTTCTGGCTTAAAGAAAGAGAGTCTTATGCGAGACTCACTTTTTC 686
DB 627 |
QY 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
DB 661 |
QY 687 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 746
DB 687 |
QY 721 TACTTGTAAATGAAGCTTTCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
DB 721 |
QY 747 TACTTGTAAATGAAGCTTTCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 806
DB 747 |
QY 781 ATTGAGCAGAGTTTTTAACAGAGCTTGGCAGTGGCTCATTTGGAATGAATTCATTT 840
DB 781 |
QY 807 ATTGAGCAGAGTTTTTAACAGAGCTTGGCAGTGGCTCATTTGGAATGAATTCATTT 866
DB 807 |
QY 841 TTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 900
DB 841 |
QY 867 TTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 926
DB 867 |
QY 901 AAGGTTTTTTCAGCAGAAAATCCTTTTGTATTTATGGAATAAATTTCTTTTAGAAGAAA 960
DB 901 |
QY 927 AAGGTTTTTTCAGCAGAAAATCCTTTTGTATTTATGGAATAAATTTCTTTTAGAAGAAA 986
DB 927 |
QY 961 ACAATTTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAGAAC 1020
DB 961 |
QY 987 ACAATTTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAGAAC 1046
DB 987 |
QY 1021 ACAGATAACGCTTTTCACTTTGGATGAGATTTT 1053
DB 1021 |
QY 1047 ACAGATAACGCTTTTCACTTTGGATGAGATTTT 1079
DB 1047 |
```

RESULT 2
HSM802033
LOCUS

HSM802033 4748 bp mRNA linear HTC 22-SEP-2004

DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE

ORGANISM

REFERENCE
AUTHORS

CONSRMT
TITLE
JOURNAL

COMMENT

FEATURES
source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

Homo sapiens mRNA; cDNA DKFp761E1312 (from clone DKFp761E1312).
AL137348
GI:6807859
HTC.

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 4748)
Ansorge,W., Krieger,S., Regiert,T., Rittmueller,C., Schwager,B.,
Mewes,H.W., Weill,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
The German cDNA Consortium

Direct Submission
Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuerberg, GERMANY

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.

This clone (DKFp761E1312) is available at the RZPD Deutsches
Resourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFp761E1312

Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.
Location/Qualifiers

1..4748

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="RZPD:DKFp761E1312"

/db_xref="taxon:9606"

/clone="DKFp761E1312"

/tissue_type="amygdala"

/clone_lib="761 (synonym: hamy2) . Vector pSport1; host

DH10B; sites NotI + SalI"

/dev_stage="adult"

/note="ribonuclease"

1..4748

/genes="DKFp761E1312"

47..1102

/genes="DKFp761E1312"

/codon_start=1

/product="hypothetical protein"

/protein_id="CAB70703.2"

/db_xref="GI:52545612"

/db_xref="GOA:Q9NTD8"

/db_xref="UniProt/TREMBL:Q9NTD8"

/translation="MGDPRPRAAGLDQDERSSTDSNESIKNEERLLKSSRRFVI
PFIOYDINWYKQAQASFWTAEEVDLSKDLPHWNKLKADKDYFISHILAFFAASDGI
VNEVLVERFSQVPEARCFYGFQILINVENHSEMYSLLDITVIRDPKKEFLFNALIE
TPYVKKADWALRWTDKRSFGVERVAFVAEVEFFVGSFAAI FWLKKRGLMPGLT
FNEALTSRDEGLHCDPACLMFOVLNKPSEERVREIIVDAVKIEQEPLEALPVGLIG
MNCILMKOYIEFVADRLVELGSKVFOAENPFDMENISLEKTNPFKEKRSYQRF
AWAETTDNVFTLDADF"

ORIGIN

Query Match 100.0%; Score 1053; DB 3; Length 4748;
Best Local Similarity 100.0%; Pred. No. 3,7e-269;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGCAGCCCGAAAGCGCGAAGCGCGCGCTGGATCAGATGAGATCATCTTCA 60
DB 47 ATGGGGCAGCCCGAAAGCGCGAAGCGCGCGCTGGATCAGATGAGATCATCTTCA 106
QY 61 GACACCAACGAAGTGAATTAAGTCAAAATGAAGACCACTCTTAAGAAAGAGTTCTCGC 120
DB 107 GACACCAACGAAGTGAATTAAGTCAAAATGAAGACCACTCTTAAGAAAGAGTTCTCGC 166
QY 121 CGGTTTGTGATCTTTTCCAAATCCAGTACCTGTATTTTGGAAAATGATTAACACAGGCACAG 180
DB 167 CGGTTTGTGATCTTTTCCAAATCCAGTACCTGTATTTTGGAAAATGATTAACACAGGCACAG 226

181 GCTTCCTCTTGGACGACGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAAACAAG 240
227 GCTTCTCTCTGACGACGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAAACAAG 286
241 CTTAAAGCAGATGAGAGTACTTCTCTCATCTTACCTTTTTCAGCCAGTGTAT 300
287 CTTAAAGCAGATGAGAGTACTTCTCTCATCTTACCTTTTTCAGCCAGTGTAT 346
301 GGAATTTGTAATGAAATTTTGGTGGAGCCCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
347 GGAATTTGTAATGAAATTTTGGTGGAGCCCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 406
361 CCCTGTTTCTATGGCTTTTCAAAATCTCTCGAGAATGTTTCACTCAGAGATGACAGTTTG 420
407 CCCTGTTTCTATGGCTTTTCAAAATCTCTCGAGAATGTTTCACTCAGAGATGACAGTTTG 466
421 CTGATAGACATTCATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTTGAA 480
467 CTGATAGACATTCATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTTGAA 526
481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTCGGATGGATAGCAGATGAAAA 540
527 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTCGGATGGATAGCAGATGAAAA 586
541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTAGAGAGGAGTTTCTTCTCAGCA 600
587 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTAGAGAGGAGTTTCTTCTCAGCA 646
601 TCTTTTGTCTATATTTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
647 TCTTTTGTCTATATTTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTCC 706
661 AATGAACATCATCAGCAGAGATGAAGATTTCTGTGACTTTGCTTGCCTGATGTTCCAA 720
707 AATGAACATCATCAGCAGAGATGAAGATTTCTGTGACTTTGCTTGCCTGATGTTCCAA 766
721 TACTTAGTAAATAAGCCTTCAGAGAAAGGTCGAGAGATCATTTGTATGCTGTCAAA 780
767 TACTTAGTAAATAAGCCTTCAGAGAAAGGTCGAGAGATCATTTGTATGCTGTCAAA 826
781 ATTGAGCAGGAGTTTAAACAGAACCTTGCAGTTGGCTCATTTGGAATGAATTCATT 840
827 ATTGAGCAGGAGTTTAAACAGAACCTTGCAGTTGGCTCATTTGGAATGAATTCATT 886
841 TTGATGAACAGTACATCAGTTTGTAGCTGACAGATTAATTGTGGAACCTTGGATTCTCA 900
887 TTGATGAACAGTACATCAGTTTGTAGCTGACAGATTAATTGTGGAACCTTGGATTCTCA 946
901 AAGTTTTTTCAGGAGAGAAATCCTTTGATTTTATGGAAGAAATTTCTTTAGAGAGAAA 960
947 AAGTTTTTTCAGGAGAGAAATCCTTTGATTTTATGGAAGAAATTTCTTTAGAGAGAAA 1006
961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGTCAGAAACC 1020
1007 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGTCAGAAACC 1066
1021 ACAGATAACGCTTTCACCTTTGGATGACAGATTTT 1053
1067 ACAGATAACGCTTTCACCTTTGGATGACAGATTTT 1099

RESULT 3
BC042948
LOCUS Homo sapiens cdna clone IMAGE:4798175, containing frame-shift errors.
DEFINITION BC042948
ACCESSION BC042948
VERSION BC042948.2 GI:34194000
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 4650)
Strausberg R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Udutin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Maman, A., Rodriguez, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickerson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 4650)
Director MGC Project.
Direct Submission
Submitted (09-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgi.nci.nih.gov>
On Aug 25, 2003 this sequence version replaced gi:27695575. ---
Contact: MGC help desk
Email: cgapsb-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabh,
Parvaneh Saeedi, JR Santos, Angeliue Schnerch, Ursula Skalska,
Duane Smal, Jeff Stott, Miranda Tsai, George Yang, Jacque
Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 70 Row: h Column: 19
This clone has the following problem: frame shifted.

FEATURES
source

1. 4650
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4798175"
/tissue_type="Brain, hypothalamus"
/clone_lib="NIH MGC_96"
/lab_host="DH10B"
/note="Vector: pBluescript"

ORIGIN

Query Match 98.9%; Score 1041; DB 3; Length 4650;
Best Local Similarity 99.9%; Pred. No. 5.8e-266;
Matches 1052; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 ATGGGCGACCGAAGCGGAGCGCGCGCGCTGATCAGATGAGATCATCTTCA 60
|||||

Db	72	ATGGGCGACCCGGAAAGCGCGGCGCTGGATCAGAGATGAGATCATCTTCA	131
Qy	61	GACACCAACGAAAGTGAATAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTCTCTCG	120
Db	132	GACACCAACGAAAGTGAATAAGTCAAAATGAAGAG-CACTCTAAGAAAGAGTCTCTCG	190
Qy	121	CGGTTTGTCACTTTTCCAAATCCAGTACCCTGTATTTGGAAAATGATPATAACAGGCACAG	180
Db	191	CGGTTTGTCACTTTTCCAAATCCAGTACCCTGTATTTGGAAAATGATPATAACAGGCACAG	250
Qy	181	GCTTCTCTTCGACAGCAGAGAGGTGCGATTCATCAAGGATCCCTCACTGGAAACAG	240
Db	251	GCTTCTCTTCGACAGCAGAGAGGTGCGATTCATCAAGGATCCCTCACTGGAAACAG	310
Qy	241	CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTTACCTTTTTCGAGCCAGTGTAT	300
Db	311	CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTTACCTTTTTCGAGCCAGTGTAT	370
Qy	301	GGAAATGTTAAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT	360
Db	371	GGAAATGTTAAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT	430
Qy	361	CGCTGTTCTATGGCTTCCAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG	420
Db	431	CGCTGTTCTATGGCTTCCAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG	490
Qy	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA	480
Db	491	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA	550
Qy	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAA	540
Db	551	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAA	610
Qy	541	TCCTACTTTTGGGAAAGAGTGTGGCTTGTCTGTAGAGAGATTTTCTTCTCAGGA	600
Db	611	TCCTACTTTTGGGAAAGAGTGTGGCTTGTCTGTAGAGAGATTTTCTTCTCAGGA	670
Qy	601	TCCTTTTGTCTATATTCCTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC	660
Db	671	TCCTTTTGTCTATATTCCTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC	730
Qy	661	AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA	720
Db	731	AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA	790
Qy	721	TACTTAGTAAATGAAGCTTTCAAGAAAGGTCAGGAGATCATTTGTTGATGCTGCAAA	780
Db	791	TACTTAGTAAATGAAGCTTTCAAGAAAGGTCAGGAGATCATTTGTTGATGCTGCAAA	850
Qy	781	ATTGAGCAGGAGTTTAAACAGAGCGCTTGCAGTTGGCTCATTTGGAATGAATTCATT	840
Db	851	ATTGAGCAGGAGTTTAAACAGAGCGCTTGCAGTTGGCTCATTTGGAATGAATTCATT	910
Qy	841	TTGATGAAACAGTACATTGAGTTTGTAGCTGCAGAGATTAATTTGTTGAACTTGTGATCTCA	900
Db	911	TTGATGAAACAGTACATTGAGTTTGTAGCTGCAGAGATTAATTTGTTGAACTTGTGATCTCA	970
Qy	901	AAGGTTTTTTCAGCAGAAAAATCCTTTTGAATTTTATGGAATAACATTTCTTTAGAGGAAA	960
Db	971	AAGGTTTTTTCAGCAGAAAAATCCTTTTGAATTTTATGGAATAACATTTCTTTAGAGGAAA	1030
Qy	961	ACAAATTTCTTTGAGNAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC	1020
Db	1031	ACAAATTTCTTTGAGNAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC	1090
Qy	1021	ACAGATAACGCTTTCACCTTTGGATGCAGATTTT	1053
Db	1091	ACAGATAACGCTTTCACCTTTGGATGCAGATTTT	1123

LOCUS	CR627376	3501 bp	mRNA	linear	HTC 22-SEP-2004			
DEFINITION	Homo sapiens mRNA; cDNA DKFZp686M05248 (from clone DKFZp686M05248).							
ACCESSION	CR627376							
VERSION	CR627376.1	GI:50949847						
KEYWORDS	HTC.							
SOURCE	Homo sapiens (human)							
ORGANISM	Homo sapiens							
REFERENCE	1 (bases 1 to 3501)							
AUTHORS	Ottenwälder,B., Obermaier,B., Deutschenbaur,S., Schaipp,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.							
CONSRSTM	The German cDNA Consortium							
TITLE	Direct Submission							
JOURNAL	Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764							
COMMENT	Neuberger, GERMANY Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix (Martinsried/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp686M05248) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp686M05248 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.							
FEATURES	Location/Qualifiers							
source	1..3501							
	/organism="Homo sapiens"							
	/mol_type="mRNA"							
	/db_xref="taxon:9606"							
	/clone="DKFZp686M05248"							
	/tissue_type="salivary gland"							
	/clone_lib="686 (synonym: hlccc3). Vector pSport1_Sfi; host DH10B; sites SfiIA + SfiIB"							
	/dev_stage="adult"							
	/notes="ribonucleotide reductase M2 B (TP53 inducible)"							
gene	1..3501							
	/gene="DKFZp686M05248"							
CDS	332..1225							
	/genes="DKFZp686M05248"							
	/codon_start=1							
	/product="hypothetical protein"							
	/protein_id="CAH10473.1"							
	/db_xref="GI:50949848"							
	/translation="MYKQAQASFWTAEVVDLSKDLPHWNKKLADKDYFISHILAFPAASDGIYNNLVERFSQVQVPEARCFYGFQILINHVSEMYSLIDITYRDPKRRFLPNALETMPYVKKKADWALRWIADRKSTFGERVAFVAAVEGVFSGSPAAIIFWLKRGFLM PGLTFSNELISRDEGLHCDPACIMFOYLNNKPPPEVRVREIIVDAVKIEQRFLEALPV GLIGMNCILMKQYIEFVADRLLVELGFSKVFOAENPFDPMENISLEKTNFFPEKRVSE YQRFVAVMAETTDNVFTLDADF"							
ORIGIN								
Query Match	96.0%; Score 1010.6; DB 3; Length 3501;							
Best Local Similarity	98.6%; Pred. No. 6.8e-258;							
Matches 1019; Conservative	0; Mismatches 14; Indels 0; Gaps 0;							
Qy	21	GGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCAGACACCAACGAAAGTGAAT	80					
Db	190	GGACGACGCTTGGAGGCGCTCAGGCGGAGATCATCTTCAGACACCAACGAAAGTGAAT	249					
Qy	81	AAAGTCAAAATGAAGAGCCACTCTCTAAGAAGAGTTCTCGCGGTTTGTCTATCTTTCCAAT	140					
Db	250	AAAGTCAAAATGAAGAGCCACTCTCTAAGAAGAGTTCTCGCGGTTTGTCTATCTTTCCAAT	309					
Qy	141	CCAGTACCTTGATATTTGGAAAATGTATAAACAGGCACAGGCTTCCTTCTGGACAGCAGA	200					
Db	310	CCAGTACCTTGATATTTGGAAAATGTATAAACAGGCACAGGCTTCCTTCTGGACAGCAGA	369					
Qy	201	AGAGGTGCGATTTCAAGAGGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTA	260					
Db	370	AGAGGTGCGATTTCAAGAGGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTA	429					

QY 261 CTTTATCTCTCACAATCTTAGCCCTTTTTCAGCCAGTGATGGAATTTGTAATGAAAATTT 320
DB 430 CTTTATCTCTCACAATCTTAGCCCTTTTTCAGCCAGTGATGGAATTTGTAATGAAAATTT 489
QY 321 GGTGGAGCCCTTTAGTCAGAGGTGTCAGAGGTTCAGAGGCTCGCTGTTTCTATGCGCTTCA 380
DB 490 GGTGGAGCCCTTTAGTCAGAGGTGTCAGAGGTTCAGAGGCTCGCTGTTTCTATGCGCTTCA 549
QY 381 AATTCTCATCGAGAATGTTTCACTCAGAGATGATACAGTTTCTGATAGACACTTACATCAG 440
DB 550 AATTCTCATCGAGAATGTTTCACTCAGAGATGATACAGTTTCTGATAGACACTTACATCAG 609
QY 441 AGATCCCAAGAAAAGGGAATTTTATTAATGCAATTTGAAACCATGCCCTATGTTAAGAA 500
DB 610 AGATCCCAAGAAAAGGGAATTTTATTAATGCAATTTGAAACCATGCCCTATGTTAAGAA 669
QY 501 AAAAGCAGATTTGGCCCTTCGATGATAGACAGATAAGAAAATCTATTTTGGGGAAGAGT 560
DB 670 AAAAGCAGATTTGGCCCTTCGATGATAGACAGATAAGAAAATCTATTTTGGGGAAGAGT 729
QY 561 GGTGGCCCTTTGCTGTAGAGAGGTTTCTTCTCAGGATCTTTTCTGCTATATTCG 820
DB 730 GGTGGCCCTTTGCTGTAGAGAGGTTTCTTCTCAGGATCTTTTCTGCTATATTCG 789
QY 621 GCTTAAAGAGAGAGGTTTATGCGAGACTCACTTTTCCAAATGAACTCATCAGCAGAGA 680
DB 790 GCTTAAAGAGAGAGGTTTATGCGAGACTCACTTTTCCAAATGAACTCATCAGCAGAGA 849
QY 681 TGAAGACTTCACTGTGACTTTTCTGCTGCTGATTTTCCAAATCTTAGTAAATAGCCTTC 740
DB 850 TGAAGACTTCACTGTGACTTTTCTGCTGCTGATTTTCCAAATCTTAGTAAATAGCCTTC 909
QY 741 AGAAGAAAGGTCAGGAGATCATTTGATGCTGTCAAAATTTGAGCAGAGTTTTTAAC 800
DB 910 AGAAGAAAGGTCAGGAGATCATTTGATGCTGTCAAAATTTGAGCAGAGTTTTTAAC 969
QY 801 AGAAGCCTTGCAGTTGGCTCATTTGGAATGAATTCATTTTGGTGAACACGATCATCGA 860
DB 970 AGAAGCCTTGCAGTTGGCTCATTTGGAATGAATTCATTTTGGTGAACACGATCATCGA 1029
QY 861 GTTTGTAGCTGACAGATTAATTTGTGAACTTGGATTTCAAGGTTTTTTCAGCAGAAAA 920
DB 1030 GTTTGTAGCTGACAGATTAATTTGTGAACTTGGATTTCAAGGTTTTTTCAGCAGAAAA 1089
QY 921 TCCTTTTGAATTTAGGAAAAATTTTCTTTTGAAGAAAAAACAATTTCTTTGAGAAACG 980
DB 1090 TCCTTTTGAATTTAGGAAAAATTTTCTTTTGAAGAAAAAACAATTTCTTTGAGAAACG 1149
QY 981 AGTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAAGAACCAAGATTAAGCTTTTCACTT 1040
DB 1150 AGTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAAGAACCAAGATTAAGCTTTTCACTT 1209
QY 1041 GGATGCAAGATTTT 1053
DB 1210 GGATGCAAGATTTT 1222

RESULT 5
AY398973
LOCUS
DEFINITION Homo sapiens HCM0069 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY398973
VERSION AY398973.1 GI:39754962
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1056)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,

Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 1056)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
Location/Qualifiers
1..1056
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1056
/locus_tag="HCM0069"

Query Match 89.7%; Score 945; DB 9; Length 1056;
Best Local Similarity 89.7%; Pred. No. 1.5e-240;
Matches 945; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
QY 1 ATGGCGGACCGGAAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 1 ATGGCGGACCGGAAAGCGCGGCTGGATCAGGATGAGAGATCATCTCA 60
QY 61 GACACCAAGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTGC 120
DB 61 GACACCAAGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTGC 120
QY 121 CGGTTTGTCTATCTTTCCAAATCCAGTACCCTGATATTTGGAATGTTAACAAGGCAAG 180
DB 121 CGGTTTGTCTATCTTTCCAAATCCAGTACCCTGATATTTGGAATGTTAACAAGGCAAG 180
QY 181 GCTTCTCTTGGACAGCAGAGAGGTGCTGATTAACAAGGATCTCCCTCACTGGAACAAG 240
DB 181 GCTTCTCTTGGACAGCAGAGAGGTGCTGATTAACAAGGATCTCCCTCACTGGAACAAG 240
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCTCACTTTAGCCCTTTTTCAGCCAGTGAT 300
DB 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCTCACTTTAGCCCTTTTTCAGCCAGTGAT 300
QY 301 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAAGGCTCAGAGGCT 360
DB 301 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAAGGCTCAGAGGCT 360
QY 361 CGCTGTTTCTATGCGCTTCAAAATTTCTCATGAGAATGTTTCACTCAGAGATGTACAGTTTG 420
DB 361 CGCTGTTTCTATGCGCTTCAAAATTTCTCATGAGAATGTTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
DB 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCTTGGCATGGATAGCAGATAGAAA 540
DB 481 NNN 540
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGTGTCTTCTCTCAGGA 600
DB 541 NNN 600
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAAAGAGAGGCTTTATGCCAGGACTCCTTTTTC 660
DB 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAAAGAGAGGCTTTATGCCAGGACTCCTTTTTC 660

[illegible]

```

/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 92"
/note="Organ: testis; Vector: pCMV-
Site 2: SalI; Cloned unidirectional
Average insert size 2.5 kb. Library
full-length clones and constructed
Note: this is a NIH MGC Library."

```

ORIGIN	Query Match Best Local Similarity 75.3%; Score 793.2; DB 5; Length 879; Matches 821; Conservative 0; Mismatches 8; Indels 3; Gaps 2;
Qy	1 ATGGGCGACCGGAAAGGCGGGAAGCGCGCGCGCTGGATCAGGATGAGAGATCATCTTCA 60
Db	
Qy	22 ATGGGCGACCGGAAAGGCGGGAAGCGCGCGCGCTGGATCAGGATGAGAGATCATCTTCA 81
Db	
Qy	61 GACACCAACGAAAGTGAATAATAAGTCAATATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Db	
Qy	82 GACACCAACGAAAGTGAATAATAAGTCAATATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 141
Db	
Qy	121 CGGTTTGTCATCTTTCCAAATCCAGTACCCCTGATATTTGGAAAAATGTATAACAGGCCAG 180
Db	
Qy	142 CGGTTTGTCATCTTTCCAAATCCAGTACCCCTGATATTTGGAAAAATGTATAACAGGCCAG 201
Db	
Qy	181 GCTTCCTTCTGGACAGCAGAGAAGGTGCACTTATCAAAAGGATCTCCCTCACTGGAACAG 240
Db	
Qy	202 GCTTCTCTTCTGGACAGCAGAGAAGGTGCACTTATCAAAAGGATCTCCCTCACTGGAACAG 261
Db	
Qy	241 CTTAAAGCAGATGAGAAGTAGTCTTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT 300
Db	
Qy	262 CTTAAAGCAGATGAGAAGTAGTCTTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT 321
Db	
Qy	301 GGAATTGTAAATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTTCCAGAGCT 360
Db	
Qy	322 GGAATTGTAAATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTTCCAGAGCT 381
Db	
Qy	361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGATGTTTCACTCAGAGATGTACAGTTTG 420
Db	
Qy	382 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGATGTTTCACTCAGAGATGTACAGTTTG 441
Db	
Qy	421 CTGATGACACTTATCATCAGAGATCCCAAGAAAGGGAAATTTTATTTAATGCAATTGAA 480
Db	
Qy	442 CTGATGACACTTATCATCAGAGATCCCAAGAAAGGGAAATTTTATTTAATGCAATTGAA 501
Db	
Qy	481 ACCATGCCCTATGTTAAAGAAAAACAGAGATTTGGGCCCTTGCGATGGATAGCAGATAGAAA 540
Db	
Qy	502 ACCATGCCCTATGTTAAAGAAAAACAGAGATTTGGGCCCTTGCGATGGATAGCAGATAGAAA 561
Db	
Qy	541 TCTACTTTTGGGAAAGAGTGGTGGCCTTTTGCTGCTGTAGAAGGAGTTTCTTCTCAGGA 600
Db	
Qy	562 TCTACTTTTGGGAAAGAGTGGTGGCCTTTTGCTGCTGTAGAAGGAGTTTCTTCTCAGGA 621
Db	
Qy	601 TCTTTTCTGCTATATTTCTGGCTAAAGAAAGAGGTCTTATGCCAGGACTCACATTTTTC 660
Db	
Qy	622 TCTTTTCTGCTATATTTCTGGCTAAAGAAAGAGGTCTTATGCCAGGACTCACATTTTTC 681
Db	
Qy	661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db	
Qy	682 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 741
Db	
Qy	721 TACTTAGTAAATAGCCCTTCAGAAAGAGGG--TCAGGGAGATCAATTTGTGATCTGTCA 778
Db	
Qy	742 TACTTAGTAAATAGCCCTTCAGAAAGAGGGTTCAGGGAGATCAATTTGTGATCTGTCA 801
Db	
Qy	779 AAATT-GAGCAGGAGTTTTTAAAGAGCCCTTGGCAGTTGGCCTCATTTGAA 829
Db	
Qy	802 AAATTGAGCAGGAGTTTTTAAAGAGCCCTTGGCAGTTGGCCTCAATTTGA 853
Db	

RESULT 9
AY398975
LOCUS

DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	

/note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH MGC Library."

REFERENCE	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524
-----------	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

RESULT 9			
AY398975			
LOCUS	1043 bp	DNA	linear
AY398975			
			GSS 12-DEC-2003

```
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATGGGCCCTTGGATGGATAGCAGATAGAA 540
Db 481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540
QY 541 TCTACTTTTGGGAAAGAGTGGGCTTTGCTGTGTAAGAGGATTTCTCTCAGGA 600
Db 541 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 600
QY 601 TCTTTTGTCTATATCTGGCTAAAGAGAGAGGCTTTATGCCAGGACTCACTTTTCC 660
Db 601 TCAATTGTGCAATATCTGGCTAAAGAGAGAGGCTCATGCTGGAGTCACTTTTCA 660
QY 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGCTTTGCTGCTGATGTTTCA 720
Db 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGCTTTGCTGCTGATGTTTCA 720
QY 721 TACTTAGTAATAAGCCTTTCAGAGAAAGGCTCAGGAGATCATTTGATGCTGTCAA 780
Db 721 TACTTAGTAATAAGCCTTTCAGAGAAAGGCTCAGGAGATCATTTGATGCTGTCAA 780
QY 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGCAAT 840
Db 781 ATCGAG-----CAGAGCCTTGGCTCATTTGGAATGCAATGCTGT 827
QY 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTTACTTTGGAACTTTGATTTCA 900
Db 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTTACTTTGGAACTTTGATTTCA 900
QY 901 AAGTTTTCAGCAGAGAAATCCTTTGATTTTATGGAAGAAATTTCTTTAGAGGAAA 960
Db 901 AAGTTTTCAGCAGAGAAATCCTTTGATTTTATGGAAGAAATTTCTTTAGAGGAAA 960
QY 961 ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGGCTTTTGCAGTTATGCGAACC 1020
Db 961 ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGGCTTTTGCAGTTATGCGAACC 1020
QY 1021 ACAGATAAAGCTTTCACCTTTGGATGAGATTT 1052
Db 1021 ACAGATAAAGCTTTCACCTTTGGATGAGATTT 1052
```

```
RESULT 10
LOCUS CN163214/c 892 bp mRNA linear EST 02-APR-2004
DEFINITION 952895 MARC 4PIG Sus scrofa cDNA 3', mRNA sequence.
ACCESSION CN163214
VERSION CN163214.1 GI:46177644
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 892)
Smith, T.P.L., Freking, B.A., Ford, J.J., Vallet, J.L., Wise, T.A.,
Noneman, D.J., Wray, J.E. and Keele, J.W.
Porcine EST collection using a normalized library constructed from
embryos representing early developmental stages
Unpublished (2003)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Place: TMR8060 row: G column: 4
Seq primer: TAGAAGGCACAGTCGAGG.
Location/Qualifiers
1..892
/organism="Sus scrofa"
/mol_type="mRNA"
```

FEATURES
source

RESULT 11
CF995079
LOCUS

CF995079 914 bp mRNA linear EST 25-NOV-2003

```
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 4PIG"
/notes="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with combined RNA from day-10, day-13,
day-15, day-25, and day-30 whole embryos."
```

ORIGIN

```
Query Match 69.4%; Score 731.2; DB 7; Length 892;
Best Local Similarity 94.1%; Pred. No. 1.4e-183;
Matches 760; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 246 AGCAGATGAGAAGTACTTCTCATCTCTCATCTTACGCTTTTTCAGCCAGTGTGGAAT 305
Db 892 ATCAGATGAGAGTATTTATCTCTCATCTTACGCTTTTTCAGCCAGTGTGGAAT 833
QY 306 TGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGTTCCAGAGGCTCGCTG 365
Db 832 TGTGAATGAAACCTTGGTGGAGCGCTTTAGTCAGGAGTGCAGTTCCAGAGGCGCTG 773
QY 366 TTTCTATGGCTTTCAAAATTTCTCATCGAATGTTCACTCAGAGATGTACAGTTTGTCTAT 425
Db 772 TTTCTACGGCTTTCAAAATTTCTCATCGAATGTTCACTCAGAGATGTACAGTTTGTCTAT 713
QY 426 AGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTAAATCAATGGAACCAT 485
Db 712 AGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTAAATCAATGGAACCAT 653
QY 486 GCCTATGTTAAGAAAAAGCAGATGGGCTTTGCGATGGATAGCAGATGAAATCTTAC 545
Db 652 GCCATATGTTAAGAAAAAGCAGATGGGCTTTGAGATGGATAGCAGATGAAAGCTTAC 593
QY 546 TTTTGGGAAAAGTGGTGGCTTTGCTGCTGTAGAAAGAGTTTCTTCTCAGAGATCTTT 605
Db 592 TTTTGGGAAAAGTGGTGGCTTTGCTGCTGTAGAAAGAGTTTCTTCTCAGAGATCTTT 533
QY 606 TGCTGCTATATCTGGCTAAAGAGAGAGCTTATGCCAGGACTCACTTTTCCAAATGA 665
Db 532 TGCTGCTATATCTGGCTAAAGAGAGAGCTTATGCCAGGACTCACTTTTCCAAATGA 473
QY 666 ACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAAATCTT 725
Db 472 ACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAGTACT 413
QY 726 AGTAAATAAGCCTTCAGAAAGAAAGGTCAGGAGATCAITTTGATGCTGTCAAAATGA 785
Db 412 GGTAAATAAGCCTTCAGAAAGAAAGGTCAGGAGATCAITTTGATGCTGTCAAAATGA 353
QY 786 GCAGGAGTTTAAACAGAGACCTTGCAGTTGGCTCATTTGGAATGAAATTCATTTTGTAT 845
Db 352 GCAGGAGTTTAAACAGAGACCTTGCAGTTGGCTCATTTGGAATGAAATTCATTTTGTAT 293
QY 846 GAAACAGTACATTGAGTTTGTAGCTGACAGATTAATTTGTGAACTTCGATTTCTCAAAGGT 905
Db 292 GAAACAGTATATTGAGTTTGTAGCTGACAGATTAATTTGTGAACTTCGATTTCTCAAAGGT 233
QY 906 TTTTTCAGGAGAAATTCCTTTTGAATTTTATGGAACAAATTTCTTTAGAGGAAACAAA 965
Db 232 TTTTTCAGGAGAAATTCCTTTTGAATTTTATGGAACAAATTTCTTTAGAGGAAACAAA 173
QY 966 TTTTCTTTGAGAAACAGATTTTCAAGATATCAGGCTTTTGGAGTTTATGCGAAGAACACAGA 1025
Db 172 TTTTCTTTGAGAAACAGATTTTCAAGATATCAGGCTTTTGGAGTTTATGCGAAGAACACAGA 113
QY 1026 TAACGCTTTCACTTGGATGCAATTTT 1053
Db 112 TAATGCTTTCACCTTAGATGCAATTTT 85
```

```

DEFINITION AGENCOURT_15621478 NIH_MGC_147 Homo sapiens cDNA clone
ACCESSION IMAGE:30528806 5', mRNA sequence.
VERSION CF95079.1 GI:38511139
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 914)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: c9apbs-r@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM615 row: f column: 15
High quality sequence stop: 677.
FEATURES
    source
        1..914
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:30528806"
            /tissue_type="Human Placenta"
            /lab_host="DH10B Tona"
            /clone_lib="NIH MGC 147"
            /note="Organ: Placenta; Vector: pBluescriptR; Site:1:
            all-XhoI; Site:2: BamH; Oligo-dT primed using primer
            5'-TTTTTTTTTTTTTTT-3', size-selected for average
            insert size 2.3 kb and normalized to 10^5. This is a
            primary library enriched for full-length clones and
            constructed using the Cap-trapper method (Carninci, in
            preparation). Library constructed by M. Brownstein
            (NIH/NHGRI, National Institutes of Health). Note: This is
            a NIH_MGC library."
ORIGIN
Query Match 69.2%; Score 728.2; DB 7; Length 914;
Best Local Similarity 99.1%; Pred. No. 9.1e-183;
Matches 764; Conservative 0; Mismatches 3; Indels 4; Gaps 3;
QY 1 ATGGGCGACCCGGAAGCGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Db 87 ATGGGCGACCCGGAAGCGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 146
QY 61 GACACCAAGCAAGTGAATTAAGTCAATGAAGACCACTCTTAAGAAGAGTTCTCC 120
Db 147 GACACCAAGCAAGTGAATTAAGTCAATGAAGACCACTCTTAAGAAGAGTTCTCC 206
QY 121 CGGTTTGTCTATCTTCCATCCAGTACCTCTGATATTGGAAATGATATAACAGGCACAG 180
Db 207 CGGTTTGTCTATCTTCCATCCAGTACCTCTGATATTGGAAATGATATAACAGGCACAG 266
QY 181 GCTTCTCTTGGACAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGACCAAG 240
Db 267 GCTTCTCTTGGACAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGACCAAG 326
QY 241 CTTAAGCAGATGAGAGTACTTCTCTCATCTTACCTTTTTCGAGCAGTGAT 300
Db 327 CTTAAGCAGATGAGAGTACTTCTCTCATCTTACCTTTTTCGAGCAGTGAT 386
QY 301 GGAATTGTAATGAAAATTGGTGGAGCGCTTTAGTCAGGAGTGCGAGGCT 360

```

```

387 GGAATTGTAATGAAAATTGGTGGAGCGCTTTAGTCAGGAGTGCGAGGCT 446
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTG 420
Db 447 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTG 506
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTATTAATGCAATTGAA 480
Db 507 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTATTAATGCAATTGAA 566
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAAAA 540
Db 567 ACCATGCCCTATGTTAAG--AAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAAAA 624
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTAGAGAGAGTCTTCTCTCAGGA 600
Db 625 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTAGAGAGAGTCTTCTCTCAGGA 684
QY 601 TCTTTTGGCTGCTATATCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACATTTTTC 660
Db 685 TCTTTTGGCTGCTATATCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACATTTTTC 744
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 720
Db 745 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 804
QY 721 TACTTAGTAAATTAAGCC-TTCAGAAAGAAAGGCTC-AGGAGATCATTTGTTG 769
Db 805 TACTTAGTAAATTAAGCCTTTCAAAAGAAAGGCTCTAGGAAGATCATTTGTTG 855

RESULT 12
LOCUS CO648380
DEFINITION ILLUMIGEN MCQ_41729 Katze MMPB2 Macaca mulatta cDNA clone
RM2B (Hs.512592), mRNA sequence.
ACCESSION CO648380.1 GI:50569874
VERSION EST.
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
REFERENCE 1 (bases 1 to 870)
AUTHORS Katze, M.G., Thomas, M., Korth, M., Iadonato, S.P. and Magnus, C.L.
TITLE Large-scale Rhesus Macaque cDNA Sequencing
JOURNAL Unpublished (2003)
COMMENT Contact: C. Magnus
Illumigen Biosciences Inc.
2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
Tel: 2063780400
Fax: 2063780408
Email: cmagnus@illumigen.com
Sequenced on 2004.06.02. 691 Q20 bases. Library Preparation: Prof.
Michael Katze Lab at University of Washington DNA Sequencing:
Illumigen Biosciences Inc. For further information, see
http://www.macaque.org
PCR Primers
FORWARD: CCTCACTAAAGGACACAAA
BACKWARD: CACTATAGGGCAATTTGGTA
Insert Length: 870 Std Error: 0.00
Plate: CL000348 row: B column: 06
Seq primer: CCCTCACTAAAGGACACAAA
POLYA=No. Location/Qualifiers
    1..870
        /organism="Macaca mulatta"
        /mol_type="mRNA"
        /strain="Indian"
        /db_xref="taxon:9544"

```

[illegible]

RESULT 13		ORIGIN	
CD656661		Query Match	64.8%; Score 682.8; DB 6; Length 804;
LOCUS		Best Local Similarity	97.3%; Pred. No. 1.1e-170;
DEFINITION		Matches 716; Conservative	0; Mismatches 17; Indels 3; Gaps 2
	804 bp mRNA		
	linear		
	EST 18-JUN-2003		
	AGENCOURT 14555875 NIA Human H1 Embryonic Stem Cell cDNA Library		


```

Db      744 CATGGACTCATCCACAGAGATGAGGAGCTTCCCTGTGAACCTTTCCTTGCCTGTATGCT 803
QY      716 TCCAATACTTAGTAAA 731
Db      804 TCCAATACTNAAGAAA 819

RESULT 15
LOCUS   BUI70979
DEFINITION BUI70979 896 bp mRNA linear EST 04-SEP-2002
5' mRNA sequence.
ACCESSION BUI70979
VERSION   BUI70979.1 GI:22684963
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 896)
AUTHORS   NIH-MGC http://mgi.nhl.nih.gov/
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-remail.nih.gov
           Tissue Procurement: ATCC
           cDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Agencourt Bioscience Corporation
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM13467 row: k column: 19
           High quality sequence stop: 619.

FEATURES             Location/Qualifiers
     1..896
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /clone="IMAGE:6143754"
         /tissue_type="retinoblastoma"
         /lab_host="DH10B (phage-resistant)"
         /clone_lib="NIH MGC 67"
         /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
         Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
         Average insert size 1.75 kb. Library constructed by Life
         Technologies."

ORIGIN
Query Match      62.2%; Score 654.8; DB 5; Length 896;
Best Local Similarity 98.7%; Pred. No. 3.4e-163;
Matches 671; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

QY      1 ATGGGCGACCCGGAAGCCGGAAGCGCGCGGTGGATCAGGATGAGAGATCATCTTCA 60
Db      55 ATGGGCGACCCGGAAGCCGGAAGCGCGCGGTGGATCAGGATGAGAGATCATCTTCA 114
QY      61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 120
Db      115 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 174
QY      121 CGTTTGTTCATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATAAAGAGGATGAT 180
Db      175 CGTTTGTTCATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATAAAGAGGATGAT 234
QY      181 GCTTCTCTTCGACAGCAGAGAGGTGCACTTATCAAGAGTCTCCCTCAGTGGAGCAAG 240
Db      235 GCTTCTCTTCGACAGCAGAGAGGTGCACTTATCAAGAGTCTCCCTCAGTGGAGCAAG 294
QY      241 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGTAT 300
Db      295 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGTAT 354

```

```

QY      301 GGAATGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
Db      355 GGAATGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 414
QY      361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db      415 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 474
QY      421 CTGATAGACACTTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db      475 CTGATAGACACTTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 534
QY      481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCCCTTGGATGATGATAGCAGATAGAAAA 540
Db      535 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCCCTTGGATGATGATAGCAGATAGAAAA 594
QY      541 TCTACTTTTGGGGAAGAGTGGTGGCCCTTGTGCTGTAGAAAGAGGTTTCTTCTCAGGA 600
Db      595 TCTACTTTTGGGGAAGAGTGGTGGCCCTTGTGCTGTAGAAAGAGGTTTCTTCTCAGGA 654
QY      601 TCTTTTGTGCTATATTTCTGGCTTAAAGAAAGAGAGGTTCTTATGCCAGGACTCACCTTTT 658
Db      655 TCTTTTGTGCTATATTTCTGGCTTAAAGAAAGAGAGGTTCTTATGCCAGGACTCACCTTTT 714
QY      659 CCAATGAACCTCATCAGCAGA 678
Db      715 CAATGAACCTCATCAGCAGA 734

```

Search completed: October 30, 2005, 03:23:14
Job time : 2726.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2005, 23:30:35 ; Search time 146 Seconds
(without alignments)
11801.366 Million cell updates/sec

Title: US-10-698-228-2
Perfect score: 1053
Sequence: 1 atggcgccaccggaaggcc.....tcacctggatgcagatttt 1053

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 81813859 residues
Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A-COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B-COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A-COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B-COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUS-COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1053	100.0	1053	4	US-10-019-733-2
2	1053	100.0	4955	4	US-10-019-733-3
3	1051.4	99.8	1053	4	US-10-019-733-12
4	1051.4	99.8	1081	4	US-10-019-733-4
5	579	55.0	2500	4	US-09-962-665-9
6	579	55.0	2500	4	US-09-963-333-9
7	579	55.0	2500	4	US-09-962-677-9
8	578.4	54.9	2479	4	US-09-949-016-2025
9	578.4	54.9	2500	4	US-09-023-655-1370
10	578.4	54.9	2500	4	US-09-949-016-145
11	441	41.9	481	3	US-08-905-223-125
12	429.8	40.8	14176	1	US-08-307-499-1
13	429.8	40.8	14176	1	US-08-307-499-14
14	429.8	40.8	14176	3	US-09-299-268-1
15	429.8	40.8	14176	3	US-09-299-268-14
16	383.2	36.4	1236	4	US-09-248-796A-3914
17	382.8	36.4	825	1	US-08-307-499-40
18	382.8	36.4	825	3	US-09-299-268-40
19	337.8	32.1	1112	1	US-08-136-743B-1
20	318.2	30.2	419	3	US-08-905-223-129
21	316	30.0	1083	4	US-09-248-796A-3915
22	260.8	24.8	656	3	US-09-328-111-791
23	195.2	18.5	449	4	US-09-270-767-14670
24	169.2	16.1	366	3	US-08-307-499-13
25	169.2	16.1	366	3	US-09-299-268-13
26	141.4	13.4	373	4	US-09-401-064-337
27	139.8	13.3	373	4	US-09-401-064-319

c	28	115.2	10.9	366	4	US-09-401-064-307	Sequence 307, Appl
	29	112.4	10.7	297	4	US-09-313-294A-4435	Sequence 4435, Ap
	30	93.8	8.9	11820	4	US-09-949-016-11887	Sequence 11887, A
	31	93.8	8.9	11826	4	US-09-949-016-13767	Sequence 13767, A
	32	92	8.7	601	4	US-09-949-016-20615	Sequence 20615, A
	33	92	8.7	601	4	US-09-949-016-69028	Sequence 69028, A
	34	74.4	7.1	276	4	US-09-313-294A-3437	Sequence 3437, Ap
	35	70.2	6.7	304	4	US-09-313-294A-7143	Sequence 7143, Ap
	36	68	6.5	276	4	US-09-313-294A-4638	Sequence 4638, Ap
	37	61	5.8	1053	4	US-09-902-540-2666	Sequence 2666, Ap
c	38	61	5.8	13706	4	US-09-902-540-1124	Sequence 1124, Ap
	39	60	5.7	601	4	US-09-949-016-69031	Sequence 69031, A
	40	59.8	5.7	1141	4	US-09-806-708B-22	Sequence 22, Appl
	41	54.4	5.2	288	4	US-09-313-294A-6108	Sequence 6108, Ap
	42	52.2	5.0	276	4	US-09-313-294A-3244	Sequence 3244, Ap
	43	50	4.7	123025	4	US-09-198-452A-1	Sequence 1, Appli
	44	50	4.7	1230230	4	US-09-438-185A-1	Sequence 1, Appli
c	45	45.4	4.3	124884	4	US-09-661-596A-76	Sequence 76, Appl

ALIGNMENTS

RESULT 1
US-10-019-733-2
; Sequence 2, Application US/10019733
; Patent No. 6882917
; GENERAL INFORMATION:

; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:

Query Match 100.0%; Score 1053; DB 4; Length 1053;
Best Local Similarity 100.0%; Pred. No. 8.7e-311;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGGCGACCCCGGAAGGCGGAGCGGCGGCGGCTGGATCAGGATGAGATCATCTTCA	60
DB	1	ATGGGCGACCCCGGAAGGCGGAGCGGCGGCGGCTGGATCAGGATGAGATCATCTTCA	60
QY	61	GACACCAACGAAAGTAAATGCAATGAAGAGCCACTCTTAAAGAAAGTCTTCG	120
DB	61	GACACCAACGAAAGTAAATGCAATGAAGAGCCACTCTTAAAGAAAGTCTTCG	120
QY	121	CGTTTGTTCATCTTCCATCCAGTACCTGATATTTGGAATGTTATTAACAGGCAG	180
DB	121	CGTTTGTTCATCTTCCATCCAGTACCTGATATTTGGAATGTTATTAACAGGCAG	180
QY	181	GCTTCTCTTGACAGCAGAGAGGTCGACTTATCAAGATCTCCTCACTGGAACAAG	240
DB	181	GCTTCTCTTGACAGCAGAGAGGTCGACTTATCAAGATCTCCTCACTGGAACAAG	240
QY	241	CTTAAAGCAGATGAGAAGTACTTCTCTCATCTTACGCTTTTTCAGCCAGTGAT	300
DB	241	CTTAAAGCAGATGAGAAGTACTTCTCTCATCTTACGCTTTTTCAGCCAGTGAT	300
QY	301	GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCGAGGCT	360

Db 301 CGAAATGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTTCAGAGGCT 360
Qy 361 CGCTGTTCTATGGCTTTCAATCTCATCGAGAAATGTTCACTCAGAGATGTACATTTG 420
Db 361 CGCTGTTCTATGGCTTTCAATCTCATCGAGAAATGTTCACTCAGAGATGTACATTTG 420
Qy 421 CTGATAGACATTTATCATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 480
Db 421 CTGATAGACATTTATCATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGGATGATGATGATGATGAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGGATGATGATGATGATGAA 540
Qy 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGATAGAGAGTTCCTTCAGGA 600
Db 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGATAGAGAGTTCCTTCAGGA 600
Qy 601 TCTTTTGTCTATATTTCTGGCTTAAAGAAAGAGGTCCTTATGCCAGGACTCCTTTTCC 660
Db 601 TCTTTTGTCTATATTTCTGGCTTAAAGAAAGAGGTCCTTATGCCAGGACTCCTTTTCC 660
Qy 661 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 661 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Qy 721 TACTTAGTAATTAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 721 TACTTAGTAATTAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Qy 781 ATTGACGAGGTTTAAACAGAAAGCCTTGCAGGTTGGCTCATTTGGAATGAATTCATT 840
Db 781 ATTGACGAGGTTTAAACAGAAAGCCTTGCAGGTTGGCTCATTTGGAATGAATTCATT 840
Qy 841 TTGATGAAAACAGTACATGAGTTTGTAGCTGACAGATTAATTTGTGGAACCTTGGATTC 900
Db 841 TTGATGAAAACAGTACATGAGTTTGTAGCTGACAGATTAATTTGTGGAACCTTGGATTC 900
Qy 901 AAGTTTTTTCAGGAGAAAATCCTTTTGTATTTATGAGAAAACATTTCTTTAGAGGAAA 960
Db 901 AAGTTTTTTCAGGAGAAAATCCTTTTGTATTTATGAGAAAACATTTCTTTAGAGGAAA 960
Qy 961 ACAAAATTTCTTGAGAAAACAGTTTCAGAGTATCAGCGTTTTCAGGTTATGGCAGAAACC 1020
Db 961 ACAAAATTTCTTGAGAAAACAGTTTTCAGAGTATCAGCGTTTTCAGGTTATGGCAGAAACC 1020

RESULT 2
US-10-019-733-3
; Sequence 3, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 3
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens

; FEATURE:
; OTHER INFORMATION:
US-10-019-733-3
Query Match 100.0%; Score 1053; DB 4; Length 4955;
Best Local Similarity 100.0%; Pred. No. 2,1e-310;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGGCGACCCGAAAGCGCGGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 245 ATGGGCGACCCGAAAGCGCGGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
Qy 61 GACACCAACGAAAGTGAATAAATCAAAATGAAGAGCACTCCTTAAGAAAGAGTTCTGC 120
Db 305 GACACCAACGAAAGTGAATAAATCAAAATGAAGAGCACTCCTTAAGAAAGAGTTCTGC 364
Qy 121 CGTTTGTCTATCTTCAATCCAGTACCTTGATATTTGGAAAATGATATAACAGGCAAG 180
Db 365 CGTTTGTCTATCTTCAATCCAGTACCTTGATATTTGGAAAATGATATAACAGGCAAG 424
Qy 181 GCTTCCTCTGACAGCAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
Db 425 GCTTCCTCTGACAGCAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAACAAG 484
Qy 241 CTTAAAGCAGATGAGAAAGTACTTCACTCTCACATCTTAGCCCTTTTTCAGCCAGTAT 300
Db 485 CTTAAAGCAGATGAGAAAGTACTTCACTCTCACATCTTAGCCCTTTTTCAGCCAGTAT 544
Qy 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAAGTTCCAGAGCT 360
Db 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAAGTTTCCAGAGCT 604
Qy 361 CGCTGTTCTATGGCTTTCAAAATTTCTCATCGAATGTTCACTCAGAGATGTACAGTTG 420
Db 605 CGCTGTTCTATGGCTTTCAAAATTTCTCATCGAATGTTCACTCAGAGATGTACAGTTG 664
Qy 421 CTGATAGACATTTATCATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 480
Db 665 CTGATAGACATTTATCATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 724
Qy 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGCATGGAATGAGATGAGAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGCATGGAATGAGATGAGAAA 784
Qy 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGTTTCTTCTCAGGA 844
Qy 601 TCTTTTGTCTATATTTCTGGCTTAAAGAAAGAGAGGTCCTTATGCCAGGACTCACCTTTTCC 660
Db 845 TCTTTTGTCTATATTTCTGGCTTAAAGAAAGAGAGGTCCTTATGCCAGGACTCACCTTTTCC 904
Qy 661 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 905 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
Qy 721 TACTTAGTAATTAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 965 TACTTAGTAATTAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 1024
Qy 781 ATTGACGAGGTTTAAACAGAAAGCCTTGCAGGTTGGCTCATTTGGAATGAATTCATT 840
Db 1025 ATTGACGAGGTTTAAACAGAAAGCCTTGCAGGTTGGCTCATTTGGAATGAATTCATT 1084
Qy 841 TTGATGAAAACAGTACATGAGTTTGTAGCTGACAGATTAATTTGTGGAACCTTGGATTTCTCA 900
Db 1085 TTGATGAAAACAGTACATGAGTTTGTAGCTGACAGATTAATTTGTGGAACCTTGGATTTCTCA 1144
Qy 901 AAGTTTTTTCAGGAGAAAATCCTTTTGTATTTATGAGAAAACATTTCTTTAGAGGAAA 960
Db 1145 AAGTTTTTTCAGGAGAAAATCCTTTTGTATTTATGAGAAAACATTTCTTTAGAGGAAA 1204
Qy 961 ACAAAATTTCTTGAGAAAACAGGTTTTCAGAGTATCAGCGTTTTCAGGTTATGGCAGAAACC 1020

Db 1205 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1264
QY 1021 ACAGATAACGCTCTTACCTTTGGATGCGAGATTTT 1053
Db 1265 ACAGATAACGCTCTTACCTTTGGATGCGAGATTTT 1297

RESULT 3

US-10-019-733-12
; Sequence 12, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-12

Query Match 99.8%; Score 1051.4; DB 4; Length 1053;
Best Local Similarity 99.9%; Pred. No. 2.7e-310;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
Db 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTCAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTCAAGAAAGAGTTCTCGC 120
QY 121 CGGTTTGTCTATCTTCCAAATCCAGTACCTGTATTTGGAAATGTTATAAACAGGCACAG 180
Db 121 CGGTTTGTCTATCTTCCAAATCCAGTACCTGTATTTGGAAATGTTATAAACAGGCACAG 180
QY 181 GCTTCCTCTGACACAGAGAGGTGCGACTTATCAAGAGTCTCCCTCACTGGAAACAAG 240
Db 181 GCTTCCTCTGACACAGAGAGGTGCGACTTATCAAGAGTCTCCCTCACTGGAAACAAG 240
QY 241 CTTAAAGCAGATGAGAAAGTACTTCTCATCTCTAGCCCTTTTTCAGCCAGTAT 300
Db 241 CTTAAAGCAGATGAGAAAGTACTTCTCATCTCTAGCCCTTTTTCAGCCAGTAT 300
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
Db 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTTTCAATTTCTCATCGAGATGTTTCACTCAGATGATACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTTCAATTTCTCATCGAGATGTTTCACTCAGATGATACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 480
QY 481 ACCATGCCCTATGTTTAAAGAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAAA 540
Db 481 ACCATGCCCTATGTTTAAAGAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAAA 540

QY 541 TCTACTTTTGGGAAAGAGTGGTGGCCTTTCTGCTGTAGAAAGGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCCTTTCTGCTGTAGAAAGGAGTTTCTTCTCAGGA 600
QY 601 TCTTTTGTCTGTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGACTCCTTTTTC 660
Db 601 TCTTTTGTCTGTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGACTCCTTTTTC 660
QY 661 AATGAACTCATCAGCAGAGATGAAGGACTTTCACGTGCTGCTTGTCTGCTGATGTTCCAA 720
Db 661 AATGAACTCATCAGCAGAGATGAAGGACTTTCACGTGCTGCTTGTCTGCTGATGTTCCAA 720
QY 721 TACTTAGTAATAAGCCTTCAGAAAGAGGCTCAGGAGAGATCATTTGTGATGCTCTCAAA 780
Db 721 TACTTAGTAATAAGCCTTCAGAAAGAGGCTCAGGAGAGATCATTTGTGATGCTCTCAAA 780
QY 781 ATTGAGCAGGAGTTTAAACAGAAAGCCTTGGCCTCATTTGGAATGAATGCAAT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAAAGCCTTGGCCTCATTTGGAATGAATGCAAT 840
QY 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGTCA 900
Db 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGTCA 900
QY 901 AAGGTTTTCAGGCAGAAATCCTTTTGTATTTTATGAAACATTTCTTTAGAAAGGAAA 960
Db 901 AAGGTTTTCAGGCAGAAATCCTTTTGTATTTTATGAAACATTTCTTTAGAAAGGAAA 960
QY 961 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
QY 1021 ACAGATAACGCTCTTCACTTTGGATGCGAGATTTT 1053
Db 1021 ACAGATAACGCTCTTCACTTTGGATGCGAGATTTT 1053

RESULT 4

US-10-019-733-4
; Sequence 4, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-4

Query Match 99.8%; Score 1051.4; DB 4; Length 1081;
Best Local Similarity 99.9%; Pred. No. 2.7e-310;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
Db 20 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATGAGATCATCTTCA 79
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTCAAGAAAGAGTTCTCGC 120
Db 80 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTCAAGAAAGAGTTCTCGC 139

```
QY 121 CGTTTGTTCATCTTCCCAATCCAGTACCTCGATATTGGAAAAATGTATAAAACAGGCACAG 180
Db |||
QY 140 CGTTTGTTCATCTTCCCAATCCAGTACCTCGATATTGGAAAAATGTATAAAACAGGCACAG 199
Db |||
QY 181 GCTTCCTCTCGACAGCAGAGAGGTGCGATCTATCAAGGATCTCCCTCACTGGAAACAAG 240
Db |||
QY 200 GCTTCCTCTCGACAGCAGAGAGGTGCGATCTATCAAGGATCTCCCTCACTGGAAACAAG 259
QY 241 CTTAAGCAGATGAGAGTACTCTCTCTCATCTCTAGCCCTTTTTCGACCCAGTGT 300
Db |||
QY 260 CTTAAGCAGATGAGAGTACTCTCTCTCATCTCTAGCCCTTTTTCGACCCAGTGT 319
QY 301 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
Db |||
QY 320 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 379
QY 361 CGCTGTTCTATGGCTTTCAAAATCTCATCGAATTTGTCATCGAGATGACAGTTTG 420
Db |||
QY 380 CGCTGTTCTATGGCTTTCAAAATCTCATCGAATTTGTCATCGAGATGACAGTTTG 439
QY 421 CTGATAGACACTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTAATGCAATTGAA 480
Db |||
QY 440 CTGATAGACACTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTAATGCAATTGAA 499
QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTCGGCTTCGCGATGGATAGCAGATAGAAA 540
Db |||
QY 500 ACCATGCCCTATGTTAAGAAAAAAGCAGATTCGGCTTCGCGATGGATAGCAGATAGAAA 559
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGGAGTTTCTCTCAGGA 600
Db |||
QY 560 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGGAGTTTCTCTCAGGA 619
QY 601 TCTTTTGTGCTATATTCTGGCTTAAGAAAGAGAGTCTTATGCCAGGACTACTTTTCC 660
Db |||
QY 620 TCTTTTGTGCTATATTCTGGCTTAAGAAAGAGAGTCTTATGCCAGGACTACTTTTCC 679
QY 661 AATGAACCTCATCAGCAGATGAGGACTTCACTGTGACTTTTGTGCTGTGCTGATGTTCCAA 720
Db |||
QY 680 AATGAACCTCATCAGCAGATGAGGACTTCACTGTGACTTTTGTGCTGTGCTGATGTTCCAA 739
QY 721 TACTTAGTAAATAGCCTTCAGAGAAAGGTCAGGAGATCAITTTGATGCTGTCAAA 780
Db |||
QY 740 TACTTAGTAAATAGCCTTCAGAGAAAGGTCAGGAGATCAITTTGATGCTGTCAAA 799
QY 781 ATTGACGAGGATTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTGCAIT 840
Db |||
QY 800 ATTGACGAGGATTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTGCAIT 859
QY 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGCGAATCTGATTTCTCA 900
Db |||
QY 860 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGCGAATCTGATTTCTCA 919
QY 901 AAGGTTTTTCAGGAGAAAAATCCTTTGATTTTATGAAAAACATTTCTTTAGAGAAAAA 960
Db |||
QY 920 AAGGTTTTTCAGGAGAAAAATCCTTTGATTTTATGAAAAACATTTCTTTAGAGAAAAA 979
QY 961 ACAAAATTTCTTTGAGAAAAAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1020
Db |||
QY 980 ACAAAATTTCTTTGAGAAAAAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1039
QY 1021 ACAGATAACGCTTTCACCTTTGGATGCAAGTTT 1053
Db |||
QY 1040 ACAGATAACGCTTTCACCTTTGGATGCAAGTTT 1072
```

RESULT 5
US-09-962-665-9
; Sequence 9, Application US/09962665
; Patent No. 6537759
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: POLYLPOLYGLUTAMATE SYNTHETASE GENE SEQUENCE

```
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING THE  
; FILE REFERENCE: 11926-015004  
; CURRENT APPLICATION NUMBER: US/09/962,665  
; CURRENT FILING DATE: 2001-09-24  
; PRIOR APPLICATION NUMBER: 09/658,659  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 09/596,033  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 09/357,743  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 09/357,024  
; PRIOR FILING DATE: 1999-07-19  
; PRIOR APPLICATION NUMBER: 60/093,484  
; PRIOR FILING DATE: 1998-07-20  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 2500  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc_feature  
; LOCATION: 128, 1464  
; OTHER INFORMATION: n = g or a  
; NAME/KEY: misc_feature  
; LOCATION: 189  
; OTHER INFORMATION: n = t or g  
; NAME/KEY: misc_feature  
; LOCATION: 524  
; OTHER INFORMATION: n = c or g  
; NAME/KEY: misc_feature  
; LOCATION: 1399  
; OTHER INFORMATION: n = t or a  
; NAME/KEY: misc_feature  
; LOCATION: 1636, 1738, 2259  
; OTHER INFORMATION: n = c or t  
US-09-962-665-9
```

```
Query Match 55.0%; Score 579; DB 4; Length 2500;  
Best Local Similarity 75.0%; Pred. No. 6,4e-166;  
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;  
  
QY 89 ATGAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTTCCAAATCCAGTACC 148  
Db |||  
QY 397 AGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTCTATCTTCCCATCCAGTACC 456  
Db |||  
QY 149 CTGATATTTGGAAAAATGTATAAACAGGCACAGGCTTCTTCTCGACAGCAGAGAGTGC 208  
Db |||  
QY 457 ATGATATCTGGCAGATGTATAGAAAGCAGAGGCTTCTTTTGGACCGCCGAGGAGTGC 516  
QY 209 ACTTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAGTACTTCACT 269  
Db |||  
QY 517 ACCTCTCNAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576  
QY 269 CTCACATCTTAGCTTTTTCGAGCCAGTGAATTTGTAATGAAATTTTGGTGAGC 328  
Db |||  
QY 577 CCATGTTCTGGCTTTCTTTTGCAGCAAGGATGGCATAGTAAATGAAATTTGGTGAGC 636  
QY 329 GCTTTAGTCAGAGGTCAGAGTTCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTTCTCA 388  
Db |||  
QY 637 GATTTAGCCAAGAGTTCAGATTACAGAGCCGCTGTTTCTATGGCTTCCAAATTTGCCA 696  
QY 389 TCGAGAAATGTTCACTCAGAGATGTAAGTTTGTGTAGACACTTATCATCAGAGATCCCA 448  
Db |||  
QY 697 TCGAAAAACATACATCTGAAATGTATAGTCTCTTATTTGACACTTACATAAAGATCCCA 756  
QY 449 AGAAAAAGGAAATTTTATTAATGCAATTTGAAACCATGCTTATGTTTAAAGAAAAAGCAG 508  
Db |||  
QY 757 AAGAAAGGGAATTTCTCTTTCAATGTCATTTGAAACGATGCTTGTGTCAAGAAAGAGCAG 816  
QY 509 ATTGGGCTTTGCGATGAGTACAGATAGAAAAATCTACTTTTGGGAAAGAGTGGTGCCCT 568  
Db |||
```

```
Db 817 ACTGGGCGCTTGGCTGGATGGGGAACAAGAGGCTACCTATGTGTAACGTGTGTAGCCT 876
Qy 569 TTGCTCTCTAGAGGAGTCTTCTCTCAGGATCTTTTGTGCTGCTATATCTGCTAAAGA 628
Db 877 TTGCTGCACTGGAAGCAATTTCTTTCGGTCTTTTGGTGCATATCTGCTCAAGA 936
Qy 629 AGAGAGCTTTATGCCAGACTCATTTTTCGAATCAATCAATCAATCAATCAATCAATCAAT 688
Db 937 AACGAGGACTGATGCCCTGACATTTTCTAATGAATTTAATGAATTTAATGAATTTAATGA 996
Qy 689 TTCACCTGCACTTTGCTGCTGATGTTTCCATCTAGTAATTAATTAATTAATTAATTAATTA 748
Db 997 TACACTGTGATTTGCTTGCCTGATGTTCAAAACACCTGCTACACAAACCATCGAGGAGA 1056
Qy 749 GGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTCAGCAGGAGTCTTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAAATTAATTAATCAATGCTTTCGGATAGAAACAGGAGTCTTCTACCTGAGGCT 1116
Qy 809 TGCCAGTTGGCCTCATTTGGAATGAATTTGATTTGATGAACAGTACATTTGATTTGTAG 868
Db 1117 TGCTGTGAAGCTCATTTGGGATGAATTTGCACTCTAATGAAGCAATACATTTGATTTGTG 1176
Qy 869 CTGACAGATTAATTTGGAAGTCTGGAATTTCTCAAGGTTTTCAGGAGAAATCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAAGTGGGTTTTCAGAAAGGTTTTCAGAGTAGAGAACCCATTTG 1236
Qy 929 ATTTATGGAAGCAATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
Db 1237 ACTTTATGAGAAATATTTCACTGGAAGGAAAGCACTAACTCTTTTGAAGAGAGTAGGCG 1296
Qy 989 AGTATCAGGCTTTTGAGTATGGCAGAAACACAGATAACGCTCTTCACTTTGGATGCAG 1048
Db 1297 AGTATCAGGAGTGGAGTGAATGTCAGTCCAAACAGAGAAATCTTTTACCTTTGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360
```

RESULT 6

```
US-09-963-333-9
; Sequence 9, Application US/09963333
; Patent No. 6664062
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE SYNTHASE GENE SEQUENCE VARIANCES
; TITLE OF INVENTION: HAVING UTILITY IN DETERMINING THE TREATMENT
; FILE OF INVENTION: OF DISEASE
; FILE REFERENCE: 11926-015002
; CURRENT APPLICATION NUMBER: US/09/963,333
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128, 1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
```

```
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc_feature
; LOCATION: 1636-1738, 2259
; OTHER INFORMATION: n = c or t
US-09-963-333-9
```

Query Match 55.0%; Score 579; DB 4; Length 2500;

Best Local Similarity 75.0%; Pred. No. 6.4e-166;

Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

```
Qy 89 ATGAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTTTCAATCCAGTACC 148
Db 397 AGGATGAGCGGCTGCTGAGAGAAAAACCCCGCGCTTTGTTCATCTTCCCATCCAGTACC 456
Qy 149 CTGATATTTGGAATAATGTATAACAGGCACAGGCTTCTTCTGACACAGCAAGAGGTCG 208
Db 457 ATGATATCTGGCAGATGTATAAGAGGCACAGGCTTCTTCTTGGACCGCCGAGGAGTTG 516
Qy 209 ACTTATCAAGAGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTACTTCACT 268
Db 517 ACCTCTCNAAGGACATTCAGCACTGGGATCCCTGAAACCCGAGGAGATATTTTATAT 576
Qy 269 CTCACATCTTAGCCTTTTTCAGCCAGTGAATTTGAAATGAAATTTGGTGGAGC 328
Db 577 CCCATGTTCTGGCTTTCTTTTTCAGCAAGCGATGGCATAGTAAATGAAATTTGGTGGAGC 636
Qy 329 GCTTTAGTACAGAGGTGAGGTTCCAGAGGCTCGCTGTTCTATGGCTTTCAAAATTTCA 388
Db 637 GATTTAGCCAAAGATTTCAAGTTACAGAACCGCTGTTTCTATGGCTTCCAAATTTGCCA 696
Qy 389 TCGAGAAATGTTCACTCAGAGATGTACAGTTTGTCTAGATAGACACTTACATCAGAGATCCA 448
Db 697 TGAAGAACATACATCTCGAAATGTATAGTCTTCTTATTTGACACTTACATAAAGATCCA 756
Qy 449 AGAAAAAGGGAATTTTATTTAATGCAATTTGAAACCACTGCCCTATGTTTAAAGAAAAAGCAG 508
Db 757 AGAAAGGGAATTTCTTCTCAATGCCATGAAACGATGCCCTTGTCTCAAGAGAGGAGCAG 816
Qy 509 ATTGGGCTTGCATGAGTAGCAGATAGAAAAATCTACTTTTGGGAAAGAGTGTGGCCT 568
Db 817 ACTGGGCTTGCCTGGAATGGGACAAAGAGGCTACCTATGGTGAACGTGTGTAGCCT 876
Qy 569 TTGCTGCTGTAGAGGAGTCTTCTCTCAGGATCTTTTGTCTGCTATATTTCTGGCTAAGA 628
Db 877 TTGCTGCACTGGAAGGCAATTTCTTTTCCGGTCTTTTTCGGTCTGATATTTCTGGCTCAAGA 936
Qy 629 AGAGAGGCTCTTATGCCAGGACTCACATTTTTCGAATGAACCTCATCAGCAGAGATCAAGGAC 688
Db 937 AACGAGACTGATGCCCTGCGCTCACATTTTCTAATGAATTTATAGCAGATAGAGGTT 996
Qy 689 TTCATGTGATTTTGTGCTGCTGATGTTCGAATTTCTAGTAAATTAAGCCTTTCAAGAGAAA 748
Db 997 TACACTGTGATTTTGTGCTGCTGATGTTCAAACACCTGGTACACAAACCACTCGAGGAGA 1056
Qy 749 GGGTCAGGAGATCATTTGATGCTGTCAAAATTCAGCAGGAGTCTTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATTAATTAATCAATGCTGTTTCGGATAGAAACAGGAGTCTTCTCAGTGGCCT 1116
Qy 809 TGCCAGTTGGCCTCATTTGGAATGAATTTGATTTTGAATGAACAGTACATTTGATTTGTAG 868
Db 1117 TGCTGTGAAGCTCATTTGGGATGAATTTGCACTCTAATGAAGCAATACATTTGATTTGTG 1176
Qy 869 CTGACAGATTAATTTGGAAGTCTTGGATTTCTCAAGGTTTTCAGGAGGAAAAATCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAAGTGGGTTTTCAGCAAGGTTTTCAGAGTAGAGAACCCATTTG 1236
Qy 929 ATTTATGGAAGCAATTTCTTTAGAGGAAACAAATTTCTTTTGAAGAAACGAGTTTCAG 988
```


Db 1237 ACTTTATGAGAATATTTCACTGGAAGGAAAGACTAACTCTTTGAGAAGAGTAGGCG 1296
QY 989 AGTATCAGCGTTTTCAGTATGAGGAGAAACACACAGATAACGCTTCCACCTTGGATGAG 1048
Db 1297 AGTATCAGAGGATGGAGTGTCAAGTCCACAGAGAATTTCTTTACCTTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360
RESULT 7
US-09-962-677-9
; Sequence 9, Application US/09962677
; Patent No. 6759200
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE PHOSPHORYLASE GENE SEQUENCE
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING
; FILE OF INVENTION: THE TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015003
; CURRENT APPLICATION NUMBER: US/09/962,677
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 128..1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc feature
; LOCATION: 1636..1738, 2259
; OTHER INFORMATION: n = c or t
US-09-962-677-9

Query Match 55.0%; Score 579; DB 4; Length 2500;
Best Local Similarity 75.0%; Pred. No. 6.4e-166;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;
QY 89 ATGAAGAGCCATCTTAAGAAGAGTTCTCGCGGTTTGTGATCTTTTCAATCCAGTACC 148
Db 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTGATCTTTCCCGCATCGAGTACC 456
QY 149 CTGATATTTGGAATATATACAGGACAGGCTTCTTCTGGACAGAGAGGTCG 208
Db 457 ATGATATCTGGCAGATGTATAGAAGGAGAGGCTTCTTTTGGACCGCGGAGGTTG 516
QY 209 ACTTATCAAGGATCTCCCTCACTGAAACAAGCTTAAAGCAGATGAGAGTACTTCACT 268
Db 517 ACCTCTCAAGGACATTCAGCACTGGGATCCCTGAAACCGGAGGAGATATTTATAT 576

QY 269 CTCACATCTTAGCCTTTTGTGACGAGTGTGAAATGTGAAATGAAATTTGGTGGAGC 328
Db 577 CCCATGTTCTGGCTTTCTTTGACGCAAGCGATGCGCATAGTAATGAAATTTGGTGGAGC 636
QY 329 GCTTTAGTCAGGAGGTGCGAGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAAATCTCA 388
Db 637 GATTTAGCCAAGAAGTTTCAGATTACAGAGCCCGCTGTTTCTATGGCTTCCAAATTGCCA 696
QY 389 TCGAGAAATGTTCACTCAGAGATGTACAGTTTGTGTAGACACTTACATCAGAGATCCCA 448
Db 697 TCGAAACACATACATTCTGAAATGTATAGTCTTCTTATTTGACACTTACATAAAGATCCCA 756
QY 449 AGAAAAGGGAATTTTATTTAATGCAATGAAACCATGCGCTATGTTTAAAGAAAAAGCAG 508
Db 757 AGAAAGGGAATTTCTTCTCAATGCCATTGAAACGATGCGCTTGTGTCAAGAAAGGCGAG 816
QY 509 ATTGGGCTTGGCATGAGATAGAGAAATCTACTTTTGGGAAAGAGTGGTGGCT 568
Db 817 ACTGGGCTTGGCTGGAATGGGACAAAGAGGCTACTATGTTGAAACGTTGTTGAGCCT 876
QY 569 TTGCTGCTGTAGAGAGGATTTTCTTCTCAGGATCTTTTCTGCTGTATATTTCTGGCTAAGA 628
Db 877 TTGCTGCACTGGAAGGCAATTTCTTTTCCGCTTCTTTGCGTCAATATTTCTGGCTCAAG 936
QY 629 AGAGAGGCTCTTATGCCAGGACTCACATTTTCCCAATGAACCTCATCAGCAGAGATGAAGGAC 688
Db 937 AAGAGGACTGATGCTGCGCTCACATTTTCTTAAGAACCTTATTTAGCAGAGATGAGGTT 996
QY 689 TTCACCTGTGACTTTGCTGCTGATGTTTCAATCTTAGTAAATAAGCCTTTCAGAAGAAA 748
Db 997 TACACTGTGATTTTGTGCTGCTGATGTTCAAAACACCTGGTACACAAACCATCGGAGAGA 1056
QY 749 GGTGAGGAGATCATTTGTTGATGCTGCTCAAAATTTGAGCAGGAGGTTTTTAAACAGAGCCT 808
Db 1057 GASTAAGAGAAATAATATCAATGCTGTTGCGGATAGAACAGGAGTTCTCTCACTGAGCCT 1116
QY 809 TGCCAGTTCGCTCATTTGGAATGAATTTGCAATTTTGTGAAACAGTACATGAGTTTGTAG 868
Db 1117 TGCTGTGAGCTCATTTGGATGAATTTGCACTCTTAATGAAGCAATATTTGAGTTTGTG 1176
QY 869 CTGACAGATTACTTGTGGAATTTGGAATTTCTCAAGGTTTTTTCAGGCGAGAAATCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAATCTGGTTTGTAGCAAGTTTTTTCAGAGTAGAGAACCAITTTG 1236
QY 929 ATTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTTCAG 988
Db 1237 ACTTTATGGAATATTTCTCACTGGAAGGAAAGACTAACTTCTTTGAGAAGAGATAGGCG 1296
QY 989 AGTATCAGGCTTTTTCAGTTATGCGAGAACACAGATACGCTTCCACCTTGGATGAGCAG 1048
Db 1297 AGTATCAGAGGATGGGATGATGTCAAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 8
US-09-949-016-2025
; Sequence 2025, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03

Query Match	54.9%;	Score 578.4;	DB 4;	Length 2500;
Best Local Similarity	75.0%;	Pred. No. 9.7e-166;		
Matches 723;	Conservative 0;	Mismatches 241;	Indels 0;	Gaps 0;
Qy	89	ATGAAGAGCCATCTCTAAGAAAGAGTTCTCGCCGGTTTGTCACTCTTTTCCAATCCAGTACC	148	
Db	397	AGGATGAGCCGCTGCTGAGAGAAACCCCGCCGCTTGTCACTTCCCATCGAGTACC	456	
Qy	149	CTGATATTTGGAAAATGTATAAACAGGCACAGCTTCTCTTCTGGACAGCAGAAGAGGTCG	208	
Db	457	ATGATATCTGGCAGATGTATAAGAGGCAGAGGCTTCTTTTGGACCCGCCGAGGAGGTTG	516	
Qy	209	ACTTATCAAAGGATCTCCCTCACTGGAACAAGCTTTAAAGCAGATGAGAAAGTACTTCATCT	268	

Db 517 ACCTCTCCAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
QY 269 CTCACATCTTAGCCTTTTTCAGCCAGTGTGAATGTAAATGAAAAATTTGGTGGAGC 328
Db 577 CCCATGTTCTGGCTTTCTTTGACGACGATGGCATAGTAATGAATAACTTTGGTGGAGC 636
QY 329 GCTTTAGTCAGAGGTGCGAGGTTCCAGAGGCTGGCTGTTTCTATGCTTTTCAAAATTCCTCA 388
Db 637 GATTTAGCCAAAGAGTTCAGATTACAGAAGCCGCTGTTTCTATGCTTCCAAATTCGCA 696
QY 389 TCGAGAAATGTTCACTCAGAGATGTACAGTTTGTCTGTATAGACATTCACATCAGAGATCCA 448
Db 697 TCGAAACATACATCTCGAAATGTATAGTCTTCTTATTCAGACATTCATATAAAGATCCA 756
QY 449 AGAAAAGGCAATTTTATTAATGCAATTAAGAACCATGCCCTATGTTAAGAAAAAGCAG 508
Db 757 AGAAAGGGAATTTCTTCAATGCAATTAAGAACCATGCCCTATGTTAAGAAAAAGCAG 816
QY 509 ATTGGGCTTGGCATGGATAGCAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCCT 568
Db 817 ACTGGGCTTGGCTGATTTGGGCAACAGAGGCTACCTATGCTGAACGTTGTAGCCT 876
QY 569 TTGCTGCTAGAGAGTTTCTTCTCAGGATCTTTTGGTGTCTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGAGTGAAGGCAATTTCTTTCCGGTCTTTTGGCTGATATTTCTGGCTCAAGA 936
QY 629 AGAGAGGCTTATGCCAGGACTCACATTTTCCAAATGAACTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGCTGATGCTGGCTCACATTTCTTAATGAACCTTATAGCAGATGAGGGT 996
QY 689 TTCACCTGTGACTTTGCTTGGCTGATTTCCAAATGAACTTAGTAATAAGCCTTCAGAAAGAA 748
Db 997 TACACTGTGATTTGCTTGGCTGATTTCAACACCTGTTACACAAACCATCGAGGAGA 1056
QY 749 GGTGAGGAGATCATTTGTGATGCTGTCAAAATTTGATGAACAGTACATGAGTTGTAG 808
Db 1057 GAGTAAAGGAATTAATATCAATGCTGTTCCGGATAGAACAGGAGTTCCTCAGTGGCCT 1116
QY 809 GGTGAGGAGATCATTTGTGATGCTGTCAAAATTTGATGAACAGTACATGAGTTGTAG 868
Db 1117 TGCCCTGAGGCTCATTTGGATGAATTCACATCTAATGAAGCAATACATGAGTTGTGG 1176
QY 869 TGCAGTGGCCTCATTTGGATGAATTTGATGAACAGTACATGAGTTGTAG 928
Db 1177 TGCCCTGAGGCTCATTTGGATGAATTTGATGAACAGTACATGAGTTGTAG 1236
QY 929 ATTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTAGAGAACGAGTTTCAG 988
Db 1237 ACTTTATGGAATATTTCACTGGAAGGAAAGACTAACTTTCTTTAGAGAGAGTAGGGG 1296
QY 989 AGTATCAGGCTTTGAGTATGCGAGAAACCAACAGATACGCTTCACCTGGATGCGAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGTCAAGTCCACAGAGAAATTTCTTTACCTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 10

US-09-949-016-145
; Sequence 145, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 145
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-145

Query Match 54.9%; Score 578.4; DB 4; Length 2500;
Best Local Similarity 75.0%; Pred. No. 9.7e-166;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCACTCTAAGAAAGAGTTCGCGCGGTTTGTCTCATCTTCCATCCAGTACC 148
Db 397 AGGATAGCGCGTGTCTGAGAGAAACCCCGCGCGTTCATCTTCCCATCGAGTACC 456
QY 149 CTGATATTTGGAAATGTATAACAGGCACAGGCTTCTCTGACAGCAAGAGAGTCG 208
Db 457 ATGATATCTGGCAGATGTATAGAGGCGAGAGGCTTCTTTTGGACCGCCGAGAGGTTG 516
QY 209 ACTTATCAAAAGGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTACTTCATCT 268
Db 517 ACCTCTCCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
QY 269 CTCACATCTTAGCCTTTTTCAGCAGCAGTGTGAATTTGTAATGAAATTTGGTGGAGC 328
Db 577 CCATGTTCTGGCTTTTTCAGCAGGAGGATAGTAAATGAAACCTTTGGTGGAGC 636
QY 329 GCTTTAGTCAGAGGTCAGAGTTCAGAGGCTCGCTGTTTCTTATGGCTTTCAAATTCCTCA 388
Db 637 GATTTAGCCAGAGTTCAGATTAAGAAGCCGCTGTTTCTATGCTTCCAAATTCGCA 696
QY 389 TCGAGAAATGTTCACTCAGAGATGTACAGTTTGTCTGTATAGACATTCACATCAGAGATCCA 448
Db 697 TCGAAACATACATCTCGAAATGTATAGTCTTCTTATGACACTTACATAAAGATGCCA 756
QY 449 AGAAAGGGAATTTTATTAATGCAATTTGAACCAATGCCCTATGTTAAGAAAAAGCAG 508
Db 757 AAGAAAGGGAATTTCTTCAATGCCATTTGAACAGATGCTTGTGTCAAGAAGAGCAG 816
QY 509 ATTGGGCTTGGCATGGATAGCAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCCT 568
Db 817 ACTGGGCTTGGCTGATTTGGGCAACAGAGGCTACCTATGTTGAACGTTGTAGCCT 876
QY 569 TTGCTGCTAGAGAGTTCCTTCTCAGGATCTTTTGGTGTCTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGAGTGAAGGCAATTTCTTTCCGGTCTTTTGGCTGATATTTCTGGCTCAAGA 936
QY 629 AGAGAGGCTTATGCCAGGACTCACATTTTCCAAATGAACTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGCTGATGCTGGCCTCACATTTTCTTAATGAACCTTATAGCAGATGAGGGT 996
QY 689 TTCACCTGTGACTTTGCTTGGCTGATTTCCAAATGAACTTAGTAATAAGCCTTCAGAAAGAA 748
Db 997 TACACTGTGATTTGCTTGGCTGATTTCAACACCTGTTACACAAACCATCGAGGAGA 1056
QY 749 GGTGAGGAGATCATTTGTGATGCTGTCAAAATTTGATGAACAGTACATGAGTTGTAG 808
Db 1057 GAGTAAAGGAATTAATATCAATGCTGTTCCGGATAGAACAGGAGTTCCTCAGTGGCCT 1116
QY 809 GGTGAGGCTCATTTGGATGAATTTGATGAACAGTACATGAGTTGTAG 868
Db 1117 TGCCCTGAGGCTCATTTGGATGAATTTGATGAACAGTACATGAGTTGTAG 1176
QY 869 CTGACAGATTTACTTGGAACTTGGATCTCAAGGTTTTTTCAGGAGGAAATCCCTTTTG 928
Db 1177 TGCCCTGAGGCTCATTTGGATGAATTTGATGAACAGTACATGAGTTGTAG 1236
QY 929 ATTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTAGAGAACGAGTTTCAG 988
Db 1237 ACTTTATGGAATATTTCACTGGAAGGAAAGACTAACTTTCTTTAGAGAGAGTAGGGG 1296
QY 989 AGTATCAGGCTTTGAGTATGCGAGAAACCAACAGATACGCTTCACCTGGATGCGAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGTCAAGTCCACAGAGAAATTTCTTTACCTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

Db 1237 ACTTATGAGAAATATTTCACTGGAGGAAAGAACTAACTCTTTTGAGAGAGAGTAGGCG 1296
QY 989 AGTATCAGCGTTTTCAGTTATGGCAGAAACACAGATAACGCTTTCACCTTGGATGCAG 1048
Db 1297 AGTATCAGAGGATGGAGTGATGTCAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 11

US-08-905-223-125
; Sequence 125, Application US/08905223
; Patent No. 622029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Win95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/905,223
FILING DATE:
CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 125:
SEQUENCE CHARACTERISTICS:
LENGTH: 481 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: CDNA
ORIGINAL SOURCE:
ORGANISM: Homo Sapiens
TISSUE TYPE: Brain
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 41..343

IDENTIFICATION METHOD: Von Heijne matrix
OTHER INFORMATION: score 4.4
OTHER INFORMATION: seq ISHILAPFAASDG/IV
US-08-905-223-125

Query Match 41.9%; Score 441; DB 3; Length 481;
Best Local Similarity 100.0%; Pred. No. 3.2e-124;
Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGACCCGGAAGCGCGGAGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Db 41 ATGGCGACCCGGAAGCGCGGAGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 100
QY 61 GACACAAACGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 101 GACACAAACGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 160

QY 121 CGTTTGTCTATCTTTTCAATCCAGTACCTGATATTTGAAATGTATAAACGACAG 180
Db 161 CGTTTGTCTATCTTTTCAATCCAGTACCTGATATTTGAAATGTATAAACGACAG 220
QY 181 GCTTCCTTTGGGACAGCAGAGAGGTGCGACTTATCAAAAGGATCTCCCTCACTGGAACAAG 240
Db 221 GCTTCCTTTGGGACAGCAGAGAGGTGCGACTTATCAAAAGGATCTCCCTCACTGGAACAAG 280
QY 241 CTTAAGACGATGAGAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGGCCAGTGAT 300
Db 281 CTTAAGACGATGAGAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGGCCAGTGAT 340
QY 301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTTAGTCAGAGGTGCGAGTTTCCAGAGGCT 360
Db 341 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTTAGTCAGAGGTGCGAGTTTCCAGAGGCT 400
QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATGTTTCACCTCAGAGATGTACAGTTTG 420
Db 401 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATGTTTCACCTCAGAGATGTACAGTTTG 460
QY 421 CTGATAGACACTTACATCAGA 441
Db 461 CTGATAGACACTTACATCAGA 481

RESULT 12

US-08-307-499-1/c
; Sequence 1, Application US/08307499
; Patent No. 5651972
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/307,499
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,241
FILING DATE: 1-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,630
FILING DATE: 29-JUN-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/342,212
FILING DATE: 21-APR-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

/ LENGTH: 14176 base pairs
 / TYPE: nucleic acid
 / STRANDEDNESS: double
 / TOPOLOGY: unknown
 / MOLECULE TYPE: DNA (genomic)
 / FEATURE:
 / NAME/KEY: CDS
 / LOCATION: 3852..4226
 / FEATURE:
 / NAME/KEY: CDS
 / LOCATION: 4585..4887
 / FEATURE:
 / NAME/KEY: CDS
 / LOCATION: 5131..5310
 / FEATURE:
 / NAME/KEY: CDS
 / LOCATION: 5760..5912
 / FEATURE:
 / NAME/KEY: CDS
 / LOCATION: 6786..7130
 / FEATURE:
 / NAME/KEY: CDS
 / LOCATION: 10148..10513
 / US-08-307-499-1

Query Match 40.8%; Score 429.8; DB 1; Length 14176;
 Best Local Similarity 66.2%; Pred. No. 5.9e-120;
 Matches 651; Conservative 0; Mismatches 327; Indels 5; Gaps 2;

QY	71	AAAGTGAATAAAGTCAATGAAGACCCACTCTTAAGAAAGAGTTCTCCCGGTTTGTC	130
Db	10841	AAAAATGATTTTATACACGAATGGAGCCATTCTTCAAGAGCTGATTCTAGGTTCTGTTA	10782
QY	131	TCCTTCCATCCAGTACCTCTGATATTTGAAATGTATAAACAAGGCACAGGCTTCCTTCT	190
Db	10781	TTTTCCCTATTAAATGATCATGATATCTGGAATGTATAAACAATCAGTGGCAAGTTTCT	10722
QY	191	GGACAGCAAGAGGTGCGACTTATCAAGAGTCTCCCTCACTGGAACAAGCTTAAAGCAG	250
Db	10721	GGACCGTGAAGAGTAGATTTTATCAAAAGATTTAGATGATGGGATAAATAAATAAG	10662
QY	251	ATGAGAAGTACTTCATCTCTCATCTTAGCCCTTTTTCAGCCAGTGTGAATGTAA	310
Db	10661	ACGAAAATACTTTATAAACAATATACTAGCATTTTTCATCTAGTGTGTTGTTAA	10602
QY	311	ATGAAAATTTGGTGAGCGCTTTAGTCAGAGGTGCGAGTTCAGAGGCTCGCTGTTCT	370
Db	10601	ATGAGAAATTTAGCGGAAAGATTTTATGTGATGTACAGTGTTCAGAGGCACGATGTTCT	10542
QY	371	ATGGCTTTCAAATCTCATCGGAATGTTCACTCAGAGATGTACAGTTTGTGATAGACA	430
Db	10541	ATGGATTTTCAAAATAGCTATGGAATAATTCATTCAGAAATGTATAGTTTATTAAATAGATA	10482
QY	431	CTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAAATGCAATTTGAAACCATGCCCT	490
Db	10481	CATATGTAAGAGTAATAATAGAAAATAATGCAATTTTAAACGCTATAGAAACAATGGAAT	10422
QY	491	ATGTTAAGAAAAGCAGATTTGGCCCTTCGGATGTAGATAGACAGATAAATCTACTTTTG	550
Db	10421	CGGTAAAAAGAAAGCTGATTTGGGCCCAAGAAATGGATATC--TAGCAACAAGGTATATG	10365
QY	551	GGAAAGAGTGTGGCCTTTGCTGCTGTAGAGAGTTTCTCTCAGGATCTTTTGCTG	610
Db	10364	GAGAAAGATGTAGCATTTTGACGCTGTGGAGGAATATTTCTTTCTGTTCAATTTGCTG	10305
QY	611	CTATATTCGGCTTAAAGAGAGAGGTTCTTATGCCAGGACTCACTTTTCCAAATGAACCTCA	670
Db	10304	CTATATTTGGATATAAACAAGAGGATGTATGCCGGATTAACATTTTCTAATGAACCTAA	10245
QY	671	TCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTCCCTGATGTTTCCAAATCTAGTAA	730
Db	10244	TAAGTAGACAGCAAGGTTTACATTTGTGATTTTGGCGGTTTAAATGTTTAAACATTTATTAC	10185

QY	731	ATAAGCCTTTCAGAAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTTGAGCAGG	790
Db	10184	ATCCACCATCTAAGGAAGTTATAAAGTCGATAATCATTTGATCGGTTAATATAGAAAAGG	10125
QY	791	AGTTTTTAACAGAAAGCCTTCCAGTTGGCTCATTTGGAATGAATTTGATTTGATGAAC	850
Db	10124	AGTTTTTGAACAGTTGCTATTCCGGTGGATCTTATAGGTATGAATTTGTTTATGTTCTC	10065
QY	851	AGTACATTCAGTTTGTAGCTGCAGAGATTACTTCTGGAACCTTGATTTCTCAAGGTTTTC	910
Db	10064	AGTATATAGAAATTCGTCGAGATAGATTATTAACAGAGTTAGGTTGTGAAAAG--TCTCA	10007
QY	911	AGGCAGAAAATCCTTTTGTATTTTATGAAAACAATTTCTTTAGAGGAAAAACAATTTCT	970
Db	10006	ATGTATATAATCCTTTTATGCTTTATGAGTATATATCACTAGAAGGTAAGACTAATTTT	9947
QY	971	TTGAGAAACAGTTTCAGAGTATCAGGTTTTCAGTTTTCAGTTATGGCAGAAACACAGATAACG	1030
Db	9946	TCGAACGACGAGTTAGTGAATATCAAAAGATGGGGTGTTTACAAAATAAAGAGAGAATA	9887
QY	1031	TCCTTCACTTGGATGCAGATTTT 1053	
Db	9886	TATTTTACGGATATAGATTTT 9864	

RESULT 13
 US-08-307-499-14
 ; Sequence 14, Application US/08307499
 ; Patent No. 5651972
 ; GENERAL INFORMATION:
 ; APPLICANT: Moyer, Richard W.
 ; APPLICANT: Vi uela, Eladio
 ; APPLICANT: Gibbs, E.P.J.
 ; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
 ; NUMBER OF SEQUENCES: 60
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: David R. Saliwanchik
 ; STREET: 2421 N.W. 41st Street, Suite A-1
 ; CITY: Gainesville
 ; STATE: Florida
 ; COUNTRY: U.S.A.
 ; ZIP: 32606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/307,499
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/908,241
 ; FILING DATE: 1-JUL-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/908,630
 ; FILING DATE: 29-JUN-1992
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/342,212
 ; FILING DATE: 21-APR-1992
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Saliwanchik, David R.
 ; REGISTRATION NUMBER: 31,794
 ; REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 904-375-8100
 ; TELEFAX: 904-372-5800
 ; INFORMATION FOR SEQ ID NO: 14:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 14176 base pairs

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 138..1460
FEATURE:
NAME/KEY: CDS
LOCATION: 2456..2659
FEATURE:
NAME/KEY: CDS
LOCATION: 2809..3030
FEATURE:
NAME/KEY: CDS
LOCATION: 3070..3330
FEATURE:
NAME/KEY: CDS
LOCATION: 3356..4180
FEATURE:
NAME/KEY: CDS
LOCATION: 4392..5894
FEATURE:
NAME/KEY: CDS
LOCATION: 6171..6398
FEATURE:
NAME/KEY: CDS
LOCATION: 6447..6875
FEATURE:
NAME/KEY: CDS
LOCATION: 6928..7431
FEATURE:
NAME/KEY: CDS
LOCATION: 7454..7858
FEATURE:
NAME/KEY: CDS
LOCATION: 7895..8155
FEATURE:
NAME/KEY: CDS
LOCATION: 8215..8682
FEATURE:
NAME/KEY: CDS
LOCATION: 8715..9539
FEATURE:
NAME/KEY: CDS
LOCATION: 9562..10272
FEATURE:
NAME/KEY: CDS
LOCATION: 10316..11908
FEATURE:
NAME/KEY: CDS
LOCATION: 11971..12780
FEATURE:
NAME/KEY: CDS
LOCATION: 12829..13107
FEATURE:
NAME/KEY: CDS
LOCATION: 13149..14171

Db 3456 GGACCGTTGAAGAAGTAGATTATCAAAAGATTAGATGATGGGATAAAATTAACATAAG 3515
Qy 251 ATGAGAAGTACTTCACTCTCAGATCTAGCCCTTTTTCAGAGCCAGTATGGAATTTAA 310
Db 3516 ACGAAAAATACTTTATATAAAACATATACTAGCATTTTTCATCTAGTATGATGTTAA 3575
Qy 311 ATGAAAAATTTGGTGAGCGCTTTAGTCAGAGGTCAGAGTTCCAGAGGCTCGCTGTTTCT 370
Db 3576 ATGAGAATTTAGCGGAAGATTTTATGTGATGTACAGTGTTCAGAGGCAGATGTTTCT 3635
Qy 371 ATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACA 430
Db 3636 ATGGAATTTCAATAGCTATGAAAAATATTCTATCAGAAATGTATAGTTTATTAATAGATA 3695
Qy 431 CTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAAACCAATGCCCT 490
Db 3696 CATATGTAAGAGATAATATAGAAAAAATGCAATTTTATTTAAGCTATAGAAAAAATGGAAT 3755
Qy 491 ATGTTAAGAAAAAGCAGATTGGGCCCTTGGGATGGATAGCAGATAGAAAAATCTACTTTTG 550
Db 3756 GCGTAAAAAGAAAGCTGATTGGGCCAGAAAAATGGATATC---TAGCAACAAGGTATATG 3812
Qy 551 GGGAAAGAGTGGTGCGCTTTGCTGTAGAAAGAGTTTCTCTCAGGATCTTTTGTCTG 610
Db 3813 GAGAAAGAGTGTAGTATGAGCTGTGAGGGAATATTCTTTTCTGTTTCAATTTGCTG 3872
Qy 611 CTATATTTCTGGCTAAAGAGAGAGGCTTTATGCGAGGACTCACATTTTCCAAATGAACTCA 670
Db 3873 CTATATTTTGGATAAAAAAAGAGGATGTATGCGCGATTAAACATTTTCTAATGAACTAA 3932
Qy 671 TCAGCAGAGATGAAGCACTTCACTGTGACTTTGCTTGCCTGATGTTCCATACCTTAGTAA 730
Db 3933 TAAGTAGAGACGAAGGTTTACATTTGTGATTTTGGCTGTTTAAATGTTTAAACATTTATTAC 3992
Qy 731 ATAAGCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCATAAATTCAGCAGG 790
Db 3993 ATCCACCATCTAAGGAAGTTATTAACGTCGATATCATTTGATCGGTTAATATAGAAAGG 4052
Qy 791 AGTTTTTAAAGAGCCTTGCAGTTGGCCTCATTTGGAATGAAATGCAATTTGATGAAAC 850
Db 4053 AGTTTTTGACAGTTGCTATTTCGGTGGATCTTATAGATGAAATTTGTTTAAATGCTC 4112
Qy 851 AGTACATTTGATTTGTAGCTGACAGATTTCTTGTGGAATTTGGAATTTCAAGGTTTTTC 910
Db 4113 AGTATATAGAATTCGTCGAGATAGATTATTAACAGAGTTAGGTTTGTGAAAG--TCTCA 4170
Qy 911 AGCAGAAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGGAAAAACAAATTTCT 970
Db 4171 ATGTATATATCTTTTAGCTTTATGGAGTATATATCCTAGAGGTAAGGTAAGTAAATTTT 4230
Qy 971 TTGAGAAACGAGTTTCAGAGATATCAGCGTTTTCAGTTTTCAGTTATGCGAGAAACACGATACG 1030
Db 4231 TCGAACGACGAGTTAGTGAATATCAAAAGATGGGGTGTGTTTACAAATAAGAGAGATA 4290
Qy 1031 TCTTCACTTGGATGCAATTTT 1053
Db 4291 TATTTTCTACGGATATAGATTTT 4313

RESULT 14

US-09-299-268-1/c
; Sequence 1, Application US/09299268
; Patent No. 6217882
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik

Query Match

Best Local Similarity 40.8%; Score 429.8; DB 1; Length 14176;
Matches 651; Conservative 0; Mismatches 327; Indels 5; Gaps 2;
Qy 71 AAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCCCGGTTTGTCA 130
Db 3336 AAATGATTTTATACAGAAATGGAGCTATTCTTCAAGAGTCTGATCTAGGTTGTTA 3395
Qy 131 TCTTTCCAAATCCAGTACCTGATATTGGAATATATAAAGCAGGACAGGCTTCTTCT 190
Db 3396 TTTTCCCTATTAAAGTATCATGATATCTGGAATATGATAAACAATCAGTGGCAAGTTT 3455
Qy 191 GGACAGCAGAGAGGTCGACTTATCAAGAGGATCTCCCTCACTGGAACAAGCTTAAGCAG 250

STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,268
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/901,127
FILING DATE:
APPLICATION NUMBER: US 07/908,241
FILING DATE: 1-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,630
FILING DATE: 29-JUN-1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/342,212
FILING DATE: 21-APR-1992
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 14176 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 3852..4226
FEATURE:
NAME/KEY: CDS
LOCATION: 4585..4887
FEATURE:
NAME/KEY: CDS
LOCATION: 5131..5310
FEATURE:
NAME/KEY: CDS
LOCATION: 5760..5912
FEATURE:
NAME/KEY: CDS
LOCATION: 6786..7130
FEATURE:
NAME/KEY: CDS
LOCATION: 10148..10513
US-09-299-268-1

Query Match 40.8%; Score 429.8; DB 3; Length 14176;
Best Local Similarity 66.2%; Pred. No. 5.9e-120;
Matches 651; Conservative 0; Mismatches 327; Indels 5; Gaps 2;

QY 71 AAGTGAATAAGTCAATGAAGAGCCACTCTAAGAAAGAGTTCCTCCCGGTTGTCA 130
DB 10841 AAAATGATTTTATACACGAATGGAGCCCTATTCTTCAAGAGTCTGATTTAGTTCGTTA 10782
QY 131 TCTTTCCAAATCCAGTACCTGATATTTGGAAAATGTATAAACAGGCACAGGCTTCCTCT 190
DB 10781 TTTTCCCTATTAAATCATGATATCTGGAAAATGTATAAACCAATCAGTGGCAAGTTT 10722

QY 191 GGACAGCAGAAAGAGTGCAGCTTATCAAGAGATCTCCCTCACTCGAAACAAGCTTAAAGCAG 250
DB 10721 GGACCGTTCGAAGAGTAGATTTATCAAAAGATTTAGATGATGGGATAAAATTAACATAAG 10662
QY 251 ATGAGAAAGTACTTCTCATCTCTCAATCTTAGCCTTTTTCAGCAGCAGTGAATGTAATGTA 310
DB 10661 ACGAAAAATACTTTATATAAACATATATCTAGCATTTTTCGATCTAGTGTGATTTGTA 10602
QY 311 ATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGCGCTCGCTGTTTCT 370
DB 10601 ATGAGAAATTTAGCGAAAGATTTATGTGATGTACAGTGTTCAGAGCGCAGATGTTTCT 10542
QY 371 ATGGCTTTCAAAATTTCTCATCGAAATGTTCACTCAGAGATGTACAGTGTTCGCTGATAGACA 430
DB 10541 ATGGATTTCAAAATAGCTATGGAATAATTCATTCAGAAATGTATAGTTTATTAATAGATA 10482
QY 431 CTTACATCAGAGATCCCAAGAAAGGAATTTTATTAATGCAATTTGAAACCAACCTGCCT 490
DB 10481 CATATGTAAGAGATAATATAGAAAAAATGCAATTTATTTAACGCTATAGAAAAACAATGGA 10422
QY 491 ATGTTAAGAAAAAGCAGATTTGGCGCTTGGATGGATAGCAGATAGAAAAATCTACTTTTG 550
DB 10421 GCGTAAAAAGAAAGCTGATTTGGCCAGAAAAATGGATATC---TAGCAACAAGTATATG 10365
QY 551 GGGAAAGAGTGGTGGCGCTTTTGCCTGTGTAGAAAGAGTTCCTTCTCAGAGATCTTTTCTG 610
DB 10364 GAGAAAGAGTGTAGCATTTGCAGCTGTGGAGGGAATATCTTTTCTGTTTCACTTTGCTG 10305
QY 611 CTATATTTCTGGCTAAAGAGAGAGGCTTTATGCGCAGGACTCACTTTTTCGAATGAACTCA 670
DB 10304 CTATATTTTGGATAAAAAAGCAGGATTTGATGCCGGAATTAACATTTTCTAATGAACATA 10245
QY 671 TCAGCAGAGATGAAGGACTTCACTGTGACTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 730
DB 10244 TAAGTAGAGACGAAGGTTTACATTTGCGTGTGTTTTCGCTGCTGCTGCTGCTGCTGCT 10185
QY 731 ATAAGCCTTCAGAAAGAAAGGTCAGGAGATCAATGTGTGATGCTGTCAAAATTCAGCAGG 790
DB 10184 ATCCACCATTAAGGAAGTTATTAACCGTCGATATCATGATGCGGTAAATATAGAAAAG 10125
QY 791 AGTTTTTAACAGAAAGCCTTGCAGTTGGCTCATATGGAATGAAATGCAATTTGATGAAAC 850
DB 10124 AGTTTTTACAGTGTGCTATTCCGCTGATCTTATAGTATGAATTTGTTTAAATGCTC 10065
QY 851 AGTACATTCAGTTTGTAGCTGCAGAGATTTCTGTGNACTTCTGAGTCTCAAGGTTTTTC 910
DB 10064 AGTATATAGAAATTCGTCGAGATAGATTTAACAAGAGTTAGGTTGTGNAAG--TCTCA 10007
QY 911 AGGCAGAAATCCTTTTGTGATTTTATGAAAAACATTTCTTTAGAAAGAAAAACAAATTTCT 970
DB 10006 ATGTATATATCTTTTGTAGCTTTATGAGATATATATCATAGAAAGGTAGACTTAATTTT 9947
QY 971 TTGAGAAACGAGTTTCAGAGTATCAGGTTTTCAGTATTTGAGTATGAGCAAAACCAAGATAACG 1030
DB 9946 TCGAACGACGAGTTAGTGAATATCAAAAGAGTGGGGTGTTCACAAATAAAGAGAGAATA 9887
QY 1031 TCTTCACTTGGATGAGATTTT 1053
DB 9886 TATTTTCTACGGATATAGATTTT 9864

RESULT 15

US-09-299-268-14
; Sequence 14, Application US/09299268
; Patent No. 6217882
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:

ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,268
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/901,127
FILING DATE:
APPLICATION NUMBER: US 07/908,241
FILING DATE: 1-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,630
FILING DATE: 29-JUN-1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/342,212
FILING DATE: 21-APR-1992
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 14176 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 138..1460
FEATURE:
NAME/KEY: CDS
LOCATION: 2456..2659
FEATURE:
NAME/KEY: CDS
LOCATION: 2809..3030
FEATURE:
NAME/KEY: CDS
LOCATION: 3070..3330
FEATURE:
NAME/KEY: CDS
LOCATION: 3356..4180
FEATURE:
NAME/KEY: CDS
LOCATION: 4392..5894
FEATURE:
NAME/KEY: CDS
LOCATION: 6171..6398
FEATURE:
NAME/KEY: CDS
LOCATION: 6447..6875
FEATURE:
NAME/KEY: CDS
LOCATION: 6928..7431
FEATURE:
NAME/KEY: CDS
LOCATION: 7454..7858
FEATURE:

NAME/KEY: CDS
LOCATION: 7895..8155
FEATURE:
NAME/KEY: CDS
LOCATION: 8215..8682
FEATURE:
NAME/KEY: CDS
LOCATION: 8715..9539
FEATURE:
NAME/KEY: CDS
LOCATION: 9562..10272
FEATURE:
NAME/KEY: CDS
LOCATION: 10316..11908
FEATURE:
NAME/KEY: CDS
LOCATION: 11971..12780
FEATURE:
NAME/KEY: CDS
LOCATION: 12829..13107
FEATURE:
NAME/KEY: CDS
LOCATION: 13149..14171
US-09-299-268-14

Query Match 40.8%; Score 429.8; DB 3; Length 14176;
Best Local Similarity 66.2%; Pred. No. 5.9e-120;
Matches 651; Conservative 0; Mismatches 327; Indels 5; Gaps 2;

QY 71 AAAGTGAATAAAGTCAAGTGAAGACCCACTCTCTAAGAAAGAGTTCTCGCGGTTTGTCA 130
|||
Db 3336 AAAATGATTTTATACACGAATGGAGCCCTATTCTTCAAGAGTCTGATTCTTAGGTTGTTA 3395

QY 131 TCTTTCCAATCCAGTACCCCTGATATTTGGAAATGTATAAAGGACAGGCTTCCTTCT 190
|||
Db 3396 TTTTCCCTATTAAATCATGATATCTGGAATAATGTATAAACAATCATAGTGCACAGTTT 3455

QY 191 GGACAGCAGAGAGGTCGACTTATCAAGAGTCTCCCTCACTGGAACAAGCTTAAAGCAG 250
|||
Db 3456 GGACCGTTGAAGAAGTAGATTTATCAAAAGATTTAGATGATGGGATAAATACTAAAG 3515

QY 251 ATGAGAAGTACTTCACTCTCTCATCTCTAGCCCTTTTTCAGGCAGTGAATGTAATCTAA 310
|||
Db 3516 ACGAAATACCTTTATAAACAATATCTAGCATTTTTCGATCTAGTGAATGTTATGTA 3575

QY 311 ATGAAATTTGGTGGAGCGCTTTAGTCAAGAGTGCAGTTCAGAGGCTCGCTGTTTCT 370
|||
Db 3576 ATGAGAAATTTAGCGGAAGATTTATGTGGATGTACAGTGTTCAGAGGCACGATGTTCT 3635

QY 371 ATGGCTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTCTGATAGACA 430
|||
Db 3636 ATGGATTTCAAAATAGCTATGGAATAATTTCACTTCAGAAATGTATAGTTTATAATAGATA 3695

QY 431 CTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAAACCATGCGCT 490
|||
Db 3696 CATATGTAAGAGATATATAGAAAATGCAATTTATTAACGTATAGAAAACATGAAT 3755

QY 491 ATGTTAAGAAAAAGCAGATTGGCGCTTCGATGGATAGCAGATAGAAAAATCTACTTTTG 550
|||
Db 3756 GCGTAAAAAAGAAAGCTGATTGGCCAGAAAATGGATATC---TAGCAACAAGTATATG 3812

QY 551 GGGAAAGAGTGTGGCTTTGCTGTGTAGAGAGGATTTTCTTCTCAGGATCTTTTCTG 610
|||
Db 3813 GAGAAAGAGTAGTAGCATTTGCAAGTGTGGAGGAAATATCTTTTCTGTTTCAATTTCTG 3872

QY 611 CTATATCTGGCTAAAGAGAGAGGTCCTTATGCGAGACTCACTTTTCCAAATGAACCTCA 670
|||
Db 3873 CTATATTTGGATAAAAAAACGAGGATTTGATGCCCGGATTAACATTTTCTTAATGAACATA 3932

QY 671 TCAGCAGAGATGAAGGACTTCTACTGTGACTTTGCTTGCCTGTGTTTCCCAATCTTAGTAA 730
|||
Db 3933 TAAGTAGAGACGAAGGTTTACATTTGATTTTGGCGTGTAAATGTTTAAACAATTTATTAC 3992

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2005, 01:51:37 ; Search time 676.5 Seconds
(without alignments)
12855.457 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1053

Sequence: 1 atggggcaccggaagcc.....tcaccttgatgcagatttt 1053

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9784742 seqs, 4129495052 residues

Total number of hits satisfying chosen parameters: 19569484

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
- 23: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 24: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq2:*
- 25: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
- 26: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 27: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
- 28: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1053	100.0	1053	19	US-10-698-228-2
2	1053	100.0	4955	18	US-10-172-118-71
3	1053	100.0	4955	19	US-10-342-887-71
4	1053	100.0	4955	19	US-10-698-228-3
5	1051.4	99.8	1053	19	US-10-698-228-12

6	1051.4	99.8	1081	19	US-10-698-228-4
7	578.4	54.9	1983	9	US-09-925-301-505
8	578.4	54.9	2216	16	US-10-084-817-342
9	578.4	54.9	2482	24	US-10-220-335-514
10	578.4	54.9	2500	9	US-09-954-456-724
11	578.4	54.9	2500	9	US-09-954-456-1169
12	578.4	54.9	2500	9	US-09-954-456-1827
13	578.4	54.9	2500	19	US-10-641-643-1370
14	578.4	54.9	2500	21	US-10-733-878-458
15	578.4	54.9	2500	22	US-10-843-641A-3751
16	578.4	54.9	2500	22	US-10-843-641A-4196
17	578.4	54.9	2500	22	US-10-843-641A-4854
18	578.4	54.9	2500	24	US-10-756-149-713
19	570.4	54.2	2113	22	US-10-764-420-1636
20	569	54.0	2641	24	US-10-220-335-170
21	561	53.3	1328	19	US-10-403-571-75
22	545.4	51.8	977	18	US-10-264-273-790
23	514	48.8	1371	24	US-10-450-763-15278
24	463	44.0	1289	26	US-11-097-143-19181
25	447.8	42.5	186854	22	US-10-872-156-34
26	442	42.0	1146	9	US-09-822-830A-174
27	432.2	41.0	1218	16	US-10-128-714-7245
28	420.4	39.9	3945	26	US-11-097-143-19180
29	403.2	38.3	963	18	US-10-369-493-34052
30	397	37.7	1146	16	US-10-128-714-2245
31	392.6	37.3	1173	18	US-10-369-493-46252
32	392	37.2	1242	17	US-10-032-585-6537
33	386.6	36.7	1292	18	US-10-369-493-27809
34	380.6	36.1	1248	19	US-10-424-599-126843
35	380.4	36.1	1200	18	US-10-369-493-45857
36	371.2	35.3	1361	19	US-10-424-599-40471
37	365.8	34.7	1586	19	US-10-424-599-73024
38	364.8	34.6	1206	18	US-10-320-797-2145
39	361.6	34.3	957	21	US-10-653-047-43
40	351.4	33.4	629	24	US-10-450-763-6851
41	349	33.1	1314	16	US-10-128-714-1245
42	349	33.1	1450	16	US-10-128-714-6245
43	349	33.1	3314	16	US-10-128-714-245
44	349	33.1	3450	16	US-10-128-714-5245
45	328.2	31.2	560	9	US-09-864-761-12381

ALIGNMENTS

RESULT 1

US-10-698-228-2
; Sequence 2, Application US/10698228
; Publication No. US2004007253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/698, 228
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019, 733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-2

Query Match 100.0%; Score 1053; DB 19; Length 1053;

Best Local Similarity 100.0%; Pred. No. 1.1e-294; Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	ATGGCGACCCGGAAAGCCCGGAAGCGCGCGGTGGATCAGGATGAGAGATCATCTTCA	60
Db	1	ATGGCGACCCGGNAGCCCGGAAGCGCGCGGTGGATCAGGATGAGAGATCATCTTCA	60
Qy	61	GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC	120
Db	61	GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC	120
Qy	121	CGGTTTGTCACTTTTCCAATCCAGTACCCGTGATATTTGGAAAATGTATAACAGGCACAG	180
Db	121	CGGTTTGTCACTTTTCCAATCCAGTACCCGTGATATTTGGAAAATGTATAACAGGCACAG	180
Qy	181	GCTTCTCTTCTGGAACAGCAGAAGAGGTGCAGCTTATCAAAAGGATCTCCCTCACTGGAACAAG	240
Db	181	GCTTCTCTTCTGGAACAGCAGAAGAGGTGCAGCTTATCAAAAGGATCTCCCTCACTGGAACAAG	240
Qy	241	CTTAAAGCAGATGAGAAAGTACTTCACTCTCACATCTTAGCCCTTTTTCGAGCCAGTGAT	300
Db	241	CTTAAAGCAGATGAGAAAGTACTTCACTCTCACATCTTAGCCCTTTTTCGAGCCAGTGAT	300
Qy	301	GGAAATGTAAATGNAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGCTTCAGAGGCT	360
Db	301	GGAAATGTAAATGNAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGCTTCAGAGGCT	360
Qy	361	CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTCACAGTTTG	420
Db	361	CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTCACAGTTTG	420
Qy	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTTAATGCAATTTGAA	480
Db	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTTAATGCAATTTGAA	480
Qy	481	ACCATGCCCTATGTTAAAGAAAAGCAGATTTGGCCCTTGCATGGATAGCAGATAGAAAA	540
Db	481	ACCATGCCCTATGTTAAAGAAAAGCAGATTTGGCCCTTGCATGGATAGCAGATAGAAAA	540
Qy	541	TCTACTTTTGGGNAAGAGTGGCCCTTTGCTGCTGTAGAAGCAGATTTCTTCTCTCAGGA	600
Db	541	TCTACTTTTGGGNAAGAGTGGGCCCTTTGCTGCTGTAGAAGCAGATTTCTTCTCTCAGGA	600
Qy	601	TCTTTTGTCTATATTTCTGGCTAAAGAAAGAGAGGTCTTATGCCAGGACTCACTTTTTC	660
Db	601	TCTTTTGTCTATATTTCTGGCTAAAGAAAGAGAGGTCTTATGCCAGGACTCACTTTTTC	660
Qy	661	AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA	720
Db	661	AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA	720
Qy	721	TACTTAGTAATAAGCCCTCAGAAGAAAGGTCAGGGAGATCATTTGTGTGATGCTGTCAAA	780
Db	721	TACTTAGTAATAAGCCCTCAGAAGAAAGGTCAGGGAGATCATTTGTGTGATGCTGTCAAA	780
Qy	781	ATTGAGCAGGAGTTTAAACAGAGCCCTTGCAGTTGGCCCTCATTTGGAATGAATTCATT	840
Db	781	ATTGAGCAGGAGTTTAAACAGAGCCCTTGCAGTTGGCCCTCATTTGGAATGAATTCATT	840
Qy	841	TTGATGAACAGTACATTTGATTTGTAGTGCAGATATCTTTGTGGAACTTTGGATTCCTCA	900
Db	841	TTGATGAACAGTACATTTGATTTGTAGTGCAGATATCTTTGTGGAACTTTGGATTCCTCA	900
Qy	901	AAGGTTTTTTCAGGCAGAAAATCCTTTTGTATTTATGGAACAATTTCTTTAGAGGAAAA	960
Db	901	AAGGTTTTTTCAGGCAGAAAATCCTTTTGTATTTATGGAACAATTTCTTTAGAGGAAAA	960
Qy	961	ACAAATTTCTTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATATGGCAGAAACC	1020
Db	961	ACAAATTTCTTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATATGGCAGAAACC	1020
Qy	1021	ACAGATAACGTCCTTTCACCTTGGATGCAGATTTT 1053	

Db 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTGCTGTAGAGGAGTCTTCTCTCAGGA 844
Qy 601 TCTTTTGTCTGTATATTCTGGCTTAAAGAGAGAGGTCTTATGCTCCAGGACTCCTCTTTTTC 660
Db 845 TCTTTTGTCTGTATATTCTGGCTTAAAGAGAGAGGTCTTATGCTCCAGGACTCCTCTTTTTC 904
Qy 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTTGCCTGTATGTTCCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTTGCCTGTATGTTCCAA 964
Qy 721 TACTTGTATTAATAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAA 780
Db 965 TACTTGTATTAATAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAA 1024
Qy 781 ATTGAGCAGAGTCTTAAACAGAGCCTTGCAGTGGCTTCACTTGGATGAATTCATTT 840
Db 1025 ATTGAGCAGAGTCTTAAACAGAGCCTTGCAGTGGCTTCACTTGGATGAATTCATTT 1084
Qy 841 TTGATGAACAGTACATTTAGTGTGCTGACAGATTAATTTGTGAACTTGGATTTCTCA 900
Db 1085 TTGATGAACAGTACATTTAGTGTGCTGACAGATTAATTTGTGAACTTGGATTTCTCA 1144
Qy 901 AAGGTTTTTCAGGAGAGAAATCCTTTTGTATTTATGAGAAACATTTCTTTAGAGGAAAA 960
Db 1145 AAGGTTTTTCAGGAGAGAAATCCTTTTGTATTTATGAGAAACATTTCTTTAGAGGAAAA 1204
Qy 961 ACAATTTCTTTGAGAAAGCAGTTTCAGAGTATCAGCGTTTTCAGTTATGAGAGAAACC 1020
Db 1205 ACAATTTCTTTGAGAAAGCAGTTTTCAGAGTATCAGCGTTTTCAGTTATGAGAGAAACC 1264
Qy 1021 ACAGATAACGCTCTTCACTTGTGATGAGATTTT 1053
Db 1265 ACAGATAACGCTCTTCACTTGTGATGAGATTTT 1297

RESULT 3

US-10-342-887-71
; Sequence 71, Application US/10342887
; Publication NO. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-71

Query Match 100.0%; Score 1053; DB 19; Length 4955;
Best Local Similarity 100.0%; Pred. No. 2.7e-294;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGCAGCCCGGAAAGCGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 245 ATGGCGCAGCCCGGAAAGCGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304

Qy 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTCTTAAGAAAGAGTTCTCGC 120
Db 305 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTCTTAAGAAAGAGTTCTCGC 364
Qy 121 CGTTTGTCTATCTTTTCCAAATCCAGTACCTGTATATTTGGAAATATGTAATAACAGGCAAG 180
Db 365 CGTTTGTCTATCTTTTCCAAATCCAGTACCTGTATATTTGGAAATATGTAATAACAGGCAAG 424
Qy 181 GCTTCCTCTTGGACACGAGAGAGGTGCACTTATCAAGAGATCTCCCTCACTGAGAACAG 240
Db 425 GCTTCCTCTTGGACACGAGAGAGGTGCACTTATCAAGAGATCTCCCTCACTGAGAACAG 484
Qy 241 CTTTAAAGCAGATCAGAAAGTACTTCTCATCTCATCTTAGCCTTTTGTGAGCCAGGTGAT 300
Db 485 CTTTAAAGCAGATCAGAAAGTACTTCTCATCTCATCTTAGCCTTTTGTGAGCCAGGTGAT 544
Qy 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAAGGCTTCCAGAGCT 360
Db 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAAGGCTTCCAGAGCT 604
Qy 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGATGTTTCACTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGATGTTTCACTCAGAGATGTACAGTTTG 664
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTTAAATGCAATTGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTTAAATGCAATTGAA 724
Qy 481 ACCATGCCCTATGTTTAAAGAAAGAGATTTGGGCTTGGGATGAGATGAGAGATAGAAA 540
Db 725 ACCATGCCCTATGTTTAAAGAAAGAGATTTGGGCTTGGGATGAGATGAGAGATAGAAA 784
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGAGAGAGTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGAGAGAGTTTCTTCTCAGGA 844
Qy 601 TCTTTTGTCTGTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACCTTTTCC 660
Db 845 TCTTTTGTCTGTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACCTTTTCC 904
Qy 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 964
Qy 721 TACTTGTATTAATAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAA 780
Db 965 TACTTGTATTAATAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAA 1024
Qy 781 ATTGAGCAGAGTCTTAAACAGAGCCTTGCAGTGGCTTCACTTGGATGAATTCATTT 840
Db 1025 ATTGAGCAGAGTCTTAAACAGAGCCTTGCAGTGGCTTCACTTGGATGAATTCATTT 1084
Qy 841 TTGATGAACAGTACATTTAGTGTGCTGACAGATTAATTTGGAACTTGGATTTCTCA 900
Db 1085 TTGATGAACAGTACATTTAGTGTGCTGACAGATTAATTTGGAACTTGGATTTCTCA 1144
Qy 901 AAGGTTTTTCAGGAGAGAAATCCTTTTGTATTTTATGAGAAACATTTCTTTAGAGGAAAA 960
Db 1145 AAGGTTTTTCAGGAGAGAAATCCTTTTGTATTTTATGAGAAACATTTCTTTAGAGGAAAA 1204
Qy 961 ACAATTTCTTTGAGAAAGCAGTTTTCAGAGTATCAGCGTTTTCAGTTATGAGAGAAACC 1020
Db 1205 ACAATTTCTTTGAGAAAGCAGTTTTCAGAGTATCAGCGTTTTCAGTTATGAGAGAAACC 1264
Qy 1021 ACAGATAACGCTCTTCACTTGTGATGAGATTTT 1053
Db 1265 ACAGATAACGCTCTTCACTTGTGATGAGATTTT 1297

RESULT 4

US-10-698-228-3
; Sequence 3, Application US/10698228


```
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 3
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-3

Query Match      100.0%; Score 1053; DB 19; Length 4955;
Best Local Similarity 100.0%; Pred. No. 2,7e-294;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGGAAGGCGGAGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 245 ATGGGCGACCCGGGAAGGCGGAGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
QY 61 GACACCAAGGAAGTGAATTAAGTCAATGAGAGCCACTCTAGAAAGAGTCTCTCC 120
DB 305 GACACCAAGGAAGTGAATTAAGTCAATGAGAGCCACTCTAGAAAGAGTCTCTCC 364
QY 121 CGGTTTGTCTATCTTCCATCCAGTACCTCTGATATTTGGAAAATGTATAAACAGGCACAG 180
DB 365 CGGTTTGTCTATCTTCCATCCAGTACCTCTGATATTTGGAAAATGTATAAACAGGCACAG 424
QY 181 GCTTCTCTTGACACGACGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAAACAAG 240
DB 425 GCTTCTCTTGACACGACGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAAACAAG 484
QY 241 CTTAAGCAGATGAGAGTACTTCTCATCTCAGATCTTACGCTTTTTCGAGGACGAT 300
DB 485 CTTAAGCAGATGAGAGTACTTCTCATCTCAGATCTTACGCTTTTTCGAGGACGAT 544
QY 301 CGAATTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
DB 545 GGAATTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 604
QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGATGTTTCACTCAGAGATGTACAGTTG 420
DB 605 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGATGTTTCACTCAGAGATGTACAGTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATGAA 480
DB 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAGAGAGATTTGGGCTTTCGATGGATAGCAGATAGAAA 540
DB 725 ACCATGCCCTATGTTAAGAAAAGAGAGATTTGGGCTTTCGATGGATAGCAGATAGAAA 784
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTGCTGTAGAGAGAGTTTCTTCTCAGGA 600
DB 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTGCTGTAGAGAGAGTTTCTTCTCAGGA 844
QY 601 TCTTTTGTCTATATCTGGCTTAAAGAGAGAGGTTCTTATGCCAGACTCACTTTTTC 660
DB 845 TCTTTTGTCTATATCTGGCTTAAAGAGAGAGGTTCTTATGCCAGACTCACTTTTTC 904
QY 661 AATGAACATCAGCAGAGATGAGGACTTCACTGTGACTTTCCTTTCCTGATGTTCCAA 920
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 3
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-2

Query Match      99.8%; Score 1051.4; DB 19; Length 1053;
Best Local Similarity 99.8%; Pred. No. 3.1e-294;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGGAAGGCGGAGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 1 ATGGGCGACCCGGGAAGGCGGAGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACCAAGGAAGTGAATTAAGTCAATGAGAGCCACTCTTAAAGAGAGTCTCTGC 120
DB 61 GACACCAAGGAAGTGAATTAAGTCAATGAGAGCCACTCTTAAAGAGAGTCTCTGC 120
QY 121 CGGTTTGTCTATCTTCCATCCAGTACCTCTGATATTTGGAAAATGTATAAACAGGCACAG 180
DB 121 CGGTTTGTCTATCTTCCATCCAGTACCTCTGATATTTGGAAAATGTATAAACAGGCACAG 180
QY 181 GCTTCTCTTGGAAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAAACAAG 240
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-12

Query Match      99.8%; Score 1051.4; DB 19; Length 1053;
Best Local Similarity 99.8%; Pred. No. 3.1e-294;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGGAAGGCGGAGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 1 ATGGGCGACCCGGGAAGGCGGAGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACCAAGGAAGTGAATTAAGTCAATGAGAGCCACTCTTAAAGAGAGTCTCTGC 120
DB 61 GACACCAAGGAAGTGAATTAAGTCAATGAGAGCCACTCTTAAAGAGAGTCTCTGC 120
QY 121 CGGTTTGTCTATCTTCCATCCAGTACCTCTGATATTTGGAAAATGTATAAACAGGCACAG 180
DB 121 CGGTTTGTCTATCTTCCATCCAGTACCTCTGATATTTGGAAAATGTATAAACAGGCACAG 180
QY 181 GCTTCTCTTGGAAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAAACAAG 240
```

```
Db 181 GCTTCCTCTGGACGACGAGAGGTTGACTTATCAAGAGATCTCCTCACTGGAAACAG 240
Qy 241 CTTAAAGCAGATGAGAAGTACTCTATCTCTACATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAAGTACTCTATCTCTACATCTTAGCCCTTTTTCAGCCAGTGAT 300
Qy 301 GGAATTGTAATGAAATTTGGTGGAGCCCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Db 301 GGAATTGTAATGAAATTTGGTGGAGCCCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Qy 361 CCCTGTTCTATGGCTTTCAAAATCTCATCGAGATGTTCACTCAGAGATGATACAGTTTG 420
Db 361 CCCTGTTCTATGGCTTTCAAAATCTCATCGAGATGTTCACTCAGAGATGATACAGTTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAGGAGATTTGGCCCTTCGATGGATAGATAGAAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAGGAGATTTGGCCCTTCGATGGATAGATAGAAAA 540
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTGAAGAGGATTTTCTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTGAAGAGGATTTTCTCTCAGGA 600
Qy 601 TCTTTTGTGCTATATCTGGCTTAAGAGAGAGGCTTTATGCCAGGACTCACTTTTCC 660
Db 601 TCTTTTGTGCTATATCTGGCTTAAGAGAGAGGCTTTATGCCAGGACTCACTTTTCC 660
Qy 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Qy 721 TACTTAGTAAATAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGTCAAA 780
Db 721 TACTTAGTAAATAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGTCAAA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATTCATT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATTCATT 840
Qy 841 TTGATGAAACAGTACATTTAGTTGTAGCTGACAGATTACTTTGTGGAACCTTCGATTTCA 900
Db 841 TTGATGAAACAGTACATTTAGTTGTAGCTGACAGATTACTTTGTGGAACCTTCGATTTCA 900
Qy 901 AAGGTTTTTCAGGCAGAAAATCCTTTTGAATTTATGGAACAATTTCTTTAGAAGAAAA 960
Db 901 AAGGTTTTTCAGGCAGAAAATCCTTTTGAATTTATGGAACAATTTCTTTAGAAGAAAA 960
Qy 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTGGAGTTATGCGAAGAAC 1020
Db 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTGGAGTTATGCGAAGAAC 1020
Qy 1021 ACAGATAACGCTTCCACCTTGGATCGAGATTTT 1053
Db 1021 ACAGATAACGCTTCCACCTTGGATCGAGATTTT 1053
```

RESULT 6

```
US-10-698-228-4
; Sequence 4, Application US/10698228
; Publication No. US2004007253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
```

```
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-4
```

Query Match 99.8%; Score 1051.4; DB 19; Length 1081;

Best Local Similarity 99.9%; Pred. No. 3.le-294; Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 1 ATGGGCGACCCGGAAGCGGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 20 ATGGGCGACCCGGAAGCGGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 79
Qy 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCACTCCCTAAGAAAGTTCGCG 120
Db 80 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCACTCCCTAAGAAAGTTCGCG 139
Qy 121 CGSTTTGTCTATCTTCCAAATCCAGTACCTCTGATATTTGGAAAAATGTATAACAGSCACAG 180
Db 140 CGSTTTGTCTATCTTCCAAATCCAGTACCTCTGATATTTGGAAAAATGTATAACAGSCACAG 199
Qy 181 GCTTCCTCTTGGACAGCAGAGAGGTGCGACTTATCAAGGATCTCCCTCACTGGAACAAAG 240
Db 200 GCTTCCTCTTGGACAGCAGAGAGGTGCGACTTATCAAGGATCTCCCTCACTGGAACAAAG 259
Qy 241 CTTAAAGCAGATGAGAAGTACTCTCTCAATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db 260 CTTAAAGCAGATGAGAAGTACTCTCTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 319
Qy 301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGCT 360
Db 320 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGCT 379
Qy 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAATGTTCACTCAGAGATGTCAGATTTG 420
Db 380 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAATGTTCACTCAGAGATGTCAGATTTG 439
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 499
Qy 481 ACCATGCCCTATGTTAAGAAAGGAGATTTGGCCCTTCGATGGATAGCAGATAGAAAA 540
Db 500 ACCATGCCCTATGTTAAGAAAGGAGATTTGGCCCTTCGATGGATAGCAGATAGAAAA 559
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGTTTTCTTCTCAGGA 600
Db 560 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGTTTTCTTCTCAGGA 619
Qy 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCCAGGACTCACATTTTCC 660
Db 620 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCCAGGACTCACATTTTCC 679
Qy 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 720
Db 680 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 739
Qy 721 TACTTAGTAAATAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGTCAAA 780
Db 740 TACTTAGTAAATAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGTCAAA 799
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGGCTTGCCTCATTTGGAATGAATTCATT 840
Db 800 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGGCTTGCCTCATTTGGAATGAATTCATT 859
```

QY 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATCTGTGAACTTGGATTCICA 900
Db |||||
860 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATCTGTGAACTTGGATTCICA 919
QY 901 AAGGTTTTTCAGGCGAGAAATCCTTTTGATTTTATGGAATAAATCTTTTGAAGAAAA 960
Db |||||
920 AAGGTTTTTCAGGCGAGAAATCCTTTTGATTTTATGGAATAAATCTTTTGAAGAAAA 979
QY 961 ACAAAATTTCTTTGAGAAAGAGTTTCAGAGTATCAGCGTTTTCGAGTTATGCGAGAAACC 1020
Db |||||
980 ACAAAATTTCTTTGAGAAAGAGTTTCAGAGTATCAGCGTTTTCGAGTTATGCGAGAAACC 1039
QY 1021 ACAGATAAGCTCTTACCTTTGGATGACAGATTTT 1053
Db |||||
1040 ACAGATAAGCTCTTACCTTTGGATGACAGATTTT 1072

RESULT 7
US-09-925-301-505
; Sequence 505, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 505
; LENGTH: 1989
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1917)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-505

Query Match 54.9%; Score 578.4; DB 9; Length 1989;
Best Local Similarity 75.0%; Pred. No. 1.2e-156;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTTCCAAATCCAGTACC 148
Db |||||
280 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTGTCTATCTTTCCCATCGAGTACC 339
QY 149 CTGATATTTGGAAATGTATAACAGGCGACAGCTTCCTTTCTGGACAGAGAGAGGTGG 208
Db |||||
340 ATGATATCTGGCAGATGTATAAGAGGCGAGAGCTTCCTTTTGGACCGCGAGGAGGTGG 399
QY 209 ACTTATCAAGATCTCCCTCACTGCAACAGCTTAAACAGATGAGAGTACTTTCATCT 268
Db |||||
400 ACTCTCAAGGACATTCAGCACTGGGAATCCTTGAACCCGAGGAGAGATATTTATAT 459
QY 269 CTCACATCTTAGCTTTTTCAGGCGAGTGTGAATTTGAAATGAAATTTGGTGGAGC 328
Db |||||
460 CCATGTTCTGGCTTCTTTGAGCAGAGCGATGATGATGAAATGAAATTTGGTGGAGC 519
QY 329 GCTTTAGTCAGAGGTGCGAGTTTCCAGAGGCTCGCTTTCTATGCGCTTCCAAATTCICA 388
Db |||||
520 GATTTAGCCAAGAGTTTCAAGTTACAGAGCGCGCTTCTATGCGCTTCCAAATTCICA 579
QY 389 TCGAGATGTTTCACTCAGAGATGTACAGTTTGTGTATACACATTCATCAGAGATCCCA 448
Db |||||
580 TCGAAACATACATCTTGAATGTATGATCTTCTTATGACACATTCATATAAGATCCCA 639
QY 449 AGAAAGGGAATTTTATTTAATGCAATTTGAACCAATGCCCTATGTTTAAAGAAAAAGCAG 508
Db |||||

Db 640 AAGAAAGGGAATTTCTCTTCAATGCCATTTGAACAGATGCTTGTGTCAGAAAGAGCGAG 699
QY 509 ATTGGCCCTTGGATGATGATAGAGATAAATCTACTTTTGGGAAAGAGTGTGCGCT 568
Db |||||
700 ACTGGCCCTTGGCTGATTTGGGCAAAAGAGGCTACCTATGTTGAACGTTGTAGGCT 759
QY 569 TTGCTCTGTAGAGGAGATTTTCTTCTCAGGATCTTTTGTGCTGATATTTCTGGCTAAAGA 628
Db |||||
760 TTGCTGCAAGTGAAGGCAATTTCTTTCCGTTCTTTTGGCTCGATATTTCTGGCTCAAGA 819
QY 629 AGAGAGTCTTATGCCAGGACTCCTTTTTCATGAATCAATCAATCAGCAGAGATGAAGAC 688
Db |||||
820 AACGAGGACTGATGCTGCGCTCACAATTTTCTAATGAACCTTATTAGCAGAGATGAGGCT 879
QY 689 TTTCACTGTGACTTTTGTGCTGCTGATGTTCCAATCTTAGTAAATAAGCCTTCAGAAGAA 748
Db |||||
880 TACACTGTGATTTTGTGCTGCTGATGTTCAACACCTGGTACACAAACCATTCGGAGGAGA 939
QY 749 GGTCTAGGAGGATCATTTGTTGATGCTGTCAAAATTTGAGCAGGAGTTTTTAAACAGAGCCT 808
Db |||||
940 GAGTAAGAGAAATAATTAATCAATGCTGTTGCGATAGAACAGGAGTTCTCTCACTGAGCCT 999
QY 809 TGCAGTTGGCCCTCATTTGGAATGAATGCAATTTGATGAACAGTACATTCAGTTGTAG 868
Db |||||
1000 TGCCTGTGAAGCTCATTTGGGATGAATTCACCTCTAATGAAGCAATACATTCAGTTGTGG 1059
QY 869 CTGACAGATTTACTTTGGAACTTGGATCTCAAGGTTTTTTCAGGCGAGAAATCCTTTTG 928
Db |||||
1060 CAGACAGACTTATGCTGGAACTGGTTTTCAGAGGTTTTTTCAGAGTAGAACCAATTTG 1119
QY 929 ATTTTATGAAACATTTCTTTAGAAAGGAAAAACAAATTTCTTTTGAGAAACGAGTTTCAG 988
Db |||||
1120 ACTTTATGAGATATTTCTCTGAGAGGAAAGACTAATCTTTTGAAGAGAGTAGGCG 1179
QY 989 AGTATCAGCGTTTTTGCAGTTATGCGAGAAACCAACAGATAACGTTCTTCCCTTGGATGAG 1048
Db |||||
1180 AGTATCAGAGGATGGGAGTGTCAAGTCCAAACAGAGAAATTTCTTTTACCTTGGATGCTG 1239
QY 1049 ATTT 1052
Db |||||
1240 ACTT 1243

RESULT 8
US-10-084-817-342
; Sequence 342, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 342
; LENGTH: 2216
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 202239.1
US-10-084-817-342

Query Match 54.9%; Score 578.4; DB 16; Length 2216;
Best Local Similarity 75.0%; Pred. No. 1.3e-156;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

```
QY 89 ATGAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTTGTCACTTTTCCAATCCAGTACC 148
Db 284 AGGATGAGCGCTGTCTGAGAGAAACCCCGCGCTTTGTCTATCTTCCCCATCGAGTACC 343
QY 149 CTGATATTTGGAAATGTATAACAGGCACAGGCTTCTTCTGGACAGCAAGAGGTG 208
Db 344 ATGATATCTGGCAGATGTATAAGAGCAGAGGCTTCTTTTGGACCGCGGAGAGGTG 403
QY 209 ACTTATCAAGAGTCTCCCTCAGTGAACAAGCTTAAGCAGATGAGAGTACTTCTATCT 268
Db 404 ACCTCTCAAGAGCATTCAGACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 463
QY 269 CTCACATCTTAGCTTTTTCAGCCAGTATGGAATGTGAATGAAATTTGGTGGAGC 328
Db 464 CCCATGTTCTGGCTTTCTTTGACGACGATGCGATAGTAATGAACTTGGTGGAGC 523
QY 329 GCTTAGTCAGAGGTGCGAGTTCCAGAGGCTCGCTGTTTCTATGCTTTCAAATTTCTCA 388
Db 524 GATTTAGCCAAGAGTTTCAGATTACAGAAGCCGCTGTTTCTATGGCTTCCAAATTGCCA 583
QY 389 TCGAGAACTGTTCACTCAGAGATGTACAGTTTGTGTAGACACTTACATCAGAGATCCCA 448
Db 584 TGAAGAACATACATCTGGAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 643
QY 449 AGAAAGGGGAATTTTATTAATGCAATTCAAACCATGCCCTATGTATGAAGAAAAAGCAG 508
Db 644 AAGAAAGGAATTTCTTCAATGCCATTTGAACGATGCTGTGTCAAGAGAGAGCAG 703
QY 509 ATTGGCCCTTGGATGATGACAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCCT 568
Db 704 ACTGGCCCTTGGCTGATTTGGGCAAAAGAGGCTACCTATGCTGAACGTTGTAGCCT 763
QY 569 TTGCTGCTGACAGAGGTTTCTTCTCAGGATCTTTTGTGCTATATTTCTGGCTAAAGA 628
Db 764 TTGCTGCTGACAGAGGCAATTTCTTTCCGGTCTTTTGGCTCGATATTTCTGGCTCAAG 823
QY 629 AGAGAGTCTTATGCCAGGACTCACATTTTCCAAATGAACTCATCAGCAGAGATGAGGAC 688
Db 824 AAGAGGACTGATGCTGGCCTCACATTTCTAATGAACTTATTAGCAGATGAGGGTT 883
QY 689 TTCACTGTGACTTGTCTGCTGATGTTCCAAATCTTAGTAATAGCCCTTCAGAGAAA 748
Db 884 TACACTGTGATTTGTCTGCTGATGTTCAACACCTGGTACACAAACCATCGAGGAGA 943
QY 749 GGGTCAGGAGATCATTTGTGATGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 808
Db 944 GAGTAAGAGAAATAATATCAATGCTGTTCCGATAGAACAGGAGTTCCTCACTGAGGCT 1003
QY 809 TGCCAGTTGGCCTCATTTGGAATGMAATGATGTTTGTGATGAAACAGTACATTTGATG 868
Db 1004 TGCCGTGGAAGCTCATTTGGGATGAAATGCACTTAATGAAGCAATACATTTGATG 1063
QY 869 CTGACAGATTAATCTGTGAACTTGGATTTCTCAAGGTTTTCAGGAGAGAAATCCCTTTG 928
Db 1064 CAGACAGACTTATGCTGGAACTGGGTTTTCAGAGGTTTTCAGAGTAGAGAACCAATTTG 1123
QY 929 ATTTATGAGAAACATTTCTTTAGAGGAAACAAATTTCTTTAGAGAAACGAGTTTCAG 988
Db 1124 ACTTTATGAGAAATATTTCACTGGAAGGAAAGACTAACTTCTTTGAGAGAGAGTAGGCG 1183
QY 989 AGTATCAGGTTTTCGAGTTATGGCAGAAACACAGATAACTCTTCACTTGGATGAG 1048
Db 1184 AGTATCAGAGGATGGGAGTGTCAAGTCCCAACAGAGAAATTTCTTTACCTTGGATGCTG 1243
QY 1049 ATTT 1052
Db 1244 ACTT 1247
```

RESULT 9

US-10-220-335-514/c

; Sequence 514, Application US/10220335

; Publication No. US20050175607A1

```
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-041
; CURRENT APPLICATION NUMBER: US/10/220,335
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 09/664,641
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,807
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/597,707
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/577,409
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 09/515,126
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 688
; SOFTWARE: Custom
; SEQ ID NO 514
; LENGTH: 2482
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-220-335-514
```

Query Match 54.9%; Score 578.4; DB 24; Length 2482;

Best Local Similarity 75.0%; Pred. No. 1.4e-156;

Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

```
QY 89 ATGAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTTGTCACTTTTCCAATCCAGTACC 148
Db 2086 AGGATGAGCGCTGTCTGAGAGAAACCCCGCGCTTTGTCTATCTTCCCATCGAGTACC 2027
QY 149 CTGATATTTGGAAATGTATAACAGGCACAGGCTTCTTCTGGACAGCAAGAGGTG 208
Db 2026 ATGATATCTGGCAGATGTATAAGAGCAGAGGCTTCTTTTGGACCGCGGAGAGGTG 1967
QY 209 ACTTATCAAGAGTCTCCCTCAGTGAACAAGCTTAAGCAGATGAGAGTACTTCTATCT 268
Db 1966 ACCTCTCAAGAGCATTTCAAGCATCTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 1907
QY 269 CTCACATCTTAGCCTTTTTCAGCCAGTGTATGGAATGTAAATGMAAAATTTGGTGGAGC 328
Db 1906 CCATGTTCTGGCTTTCTTTGACGACGATGAGTGCATAGTAATGMAAACTTTGGTGGAGC 1847
QY 329 GCTTTAGTCAGAGGTGCAAGTTCAGAGGCTCGCTGTTCTATGCTTTTCAAAATTTCA 388
Db 1846 GATTTAGCCAAGAGTTCAGATTACAGAAGCCGCTGTTCTATGGCTTCCAAATTTGCCA 1787
QY 389 TCGAGAAATGTTCACTCAGAGATGTACAGTTTGTGTAGATAGACACTTACATCAGAGATCCCA 448
Db 1786 TGGAAACATACATCTGAAATGTATAGTCTTCTTATTTGACACTTACATAAAGATCCCA 1727
QY 449 AGAAAGGGGAATTTTATTTTAAATGCAATTCAAACCATGCCCTATGTTAAGAAAAAGCAG 508
Db 1726 AAGAAAGGGAATTTCTTCAATGCCATTTGAACGATGCTTGTGTCAAGAAAGGCGAG 1667
QY 509 ATTGGCCCTTGGATGATGACAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCCT 568
Db 1666 ACTGGCCCTTGGCTGGATTTGGGCAAAAGAGGCTACCTATGGTGAACGTTGTAGCCT 1607
QY 569 TTGCTGCTGTAGAGAGTTTTCTTCTCAGAGTCTTTTGTGCTATATTTCTGGCTAAGA 628
Db 1606 TTGCTGCTGAGGAGCAATTTCTTTTCCGGTCTTTTGGCTCGATATTTCTGGCTCAAGA 1547
QY 629 AGAGAGTCTTATGCCAGGACTCACATTTTCCAAATGAACTCATCAGCAGAGATCAAGGAC 688
Db 1546 AAGAGGACTGATGCTGGCCTCACATTTTCTAATGAATTTATAGCAGATGAGGGTT 1487
QY 689 TTCACTGTGACTTTTGTGCTGATGTTCCAATCTAGTAATTAAGCCCTTCAGAGAAA 748
Db 1486 TACACTGTGATTTTGTCTGCTGATGTTTCAACACACCTGGTACACAAACCATCGAGGAGA 1427
QY 749 GGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTTGAGCAGGAGCTTTTAAACAGAGCCT 808
```

Db 1426 GAGTAAGAGAAATAATCAATGCTGTTCCGATAGAACAGGAGTTCCTCACTGAGGCT 1367
Qy 809 TCCAGATTGGCTCATTTGAATGAATTTGATGATAAGAGTACATTTGATTTGAG 868
Db 1366 TGCCCTGTGAAGCTCATTTGGGATGAATGCACTCTAATGAAGCAATACATTTGATTTGGG 1307
Qy 869 CTGACAGATTACTTTGTGGAACTTGGAATCTCAAGGTTTTCAGGCGAGAAAATCCCTTTTG 928
Db 1306 CAGACAGACTTATGCTGGAACTGGGTTTTCAGAAAGTTTTCAGATGAGAACCAATTTG 1247
Qy 929 ATTTATGGAACAATTTCTTTTAGAAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
Db 1246 ACTTTATGAGAATAATTTCACTGGAAGGAAAGACATACTTTCTTTGAGAAGAGTAGGCG 1187
Qy 989 AGTATCAGCGTTTTCAGTATGAGGAGAAACACAGATAAGTCTTCACTTGGATGAG 1048
Db 1186 AGTATCAGAGGATGGAGTGTCAAGTCCAAACAGAGAATTTCTTTTACCTTTGGATGCTG 1127
Qy 1049 ATTT 1052
Db 1126 ACTT 1123

RESULT 10
US-09-954-456-724
; Sequence 724, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; FILE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 724
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-724

Query Match 54.9%; Score 578.4; DB 9; Length 2500;
Best Local Similarity 75.0%; Pred. No. 1.4e-156;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;
Qy 89 ATGAGAGCCACTCCTAAGAGAGTTCTCGCGGTTTGTCTATCTTTCCATCCAGTACC 148
Db 397 AGGATGAGCCGCTGTGAGAGAAAACCCCGCGCTTTGTCTATCTTTCCCATCGAGTACC 456
Qy 149 CTGATATTTGGAAAATGTATAACAGGCACAGGCTTCCTTCTGACAGCAGAGAGGTCG 208

Db 457 ATGATATTCGGCAGATGTATAGAGGCGAGAGGCTTCCTTTTGGACCGCGAGAGGTTG 516
Qy 209 ACTTATCAAGAGATCTCCCTCACTGGAAACAAGCTTAAGCAGATGAGAAGTACTTCACT 268
Db 517 ACCTCTCCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
Qy 269 CTCACATCTTAGCCTTTTTCAGCCAGTGATGGAATTTGAAATGAAAAATTTGGTGAGC 328
Db 577 CCATGCTTCTGGCTTTCTTTGCAAGCAAGCGATGCGATAGTAATGAAATCTTGGTGAGC 636
Qy 329 GCTTTAGTCAGAGGTGCGAGTTCCAGAGGCTCGCTGTTCTATGCTTTTCAATTTCTCA 388
Db 637 GATTTAGCCAAGAGTTTCAGATTACAGAGCCGCTGTTTCTATGCTTCCAAATTTGCCA 696
Qy 389 TCGAGAACTCTCACTCAGAGATGATACAGTTTGTGTAGACACTTACATCAGAGATCCCA 448
Db 697 TGAAGAACATACATCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 756
Qy 449 AGAAAGGGAAATTTTATTTAATGCAATTTGAAACCATGCCCCTATGTTAAGAAAAAGCAG 508
Db 757 AAGAAGGGAAATTTCTTCAATGCCATTTGAACGATGCTTGTGCAAGAGAGGAG 816
Qy 509 ATTGGGCTTGGCATGGATAGCAGATAGAAAAATCTACTTTTGGGAAAAGAGTGGTGCT 568
Db 817 ACTGGGCTTGGCTGGATTTGGGCAAAAGAGGCTACTATGCTGAACGTTGTGTAGCCT 876
Qy 569 TTGCTGCTAGAGAGGTTTCTTCTCAGGATCTTTGCTGCTATATCTTGGCTAAGA 628
Db 877 TTGCTGCTAGAGGAGGCAATTTCTTTCCGCTTCTTTTGGCTGATATTTCTGGCTCAAGA 936
Qy 629 AGAGAGGCTCTATGCCAGGACTCACATTTTCCAAATGAACTCATCAGCAGAGATGAAGAC 688
Db 937 AACGAGACTGATGCTGCGCTCACATTTCTAATGAACCTTATAGCAGAGATGAGGTT 996
Qy 689 TTCACGTGCACTTTGCTGCTGATGTTCCAACTCTTGAATTAAGCTTTTCAGAGAAA 748
Db 997 TACACTGTGATTTGCTGCTGATGTTCAAAACACCTGGTACACAAACCATCGAGGAGA 1056
Qy 749 GGTGAGGAGATCATTTGATGCTGCTCAAAATGAGCAGGAGTTTTCACAGAGCCT 808
Db 1057 GAGTAAGAGAAATAATTAATGCTGTTCCGATAGAACAGGAGTTCTCCTCAGGCT 1116
Qy 809 TGCCAGTTGGCCTCATTTGGAATGAAATGCAATTTTGAATGAAACAGTACATTTGAGTTCTAG 868
Db 1117 TGCTGTGAGCTCATTTGGATGAATTTGCACTCTAATGAGCAATACATTTGAGTTTGG 1176
Qy 869 CTGACAGATTACTTTGTGGAACTTTGGATTTCTCAAAGTTTTCAGGCGAGAAAATCCTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACCTGGGTTTTCAGCAAGGTTTTCAGAGTAGAGAACCAATTG 1236
Qy 929 ATTTATGGAACAATTTCTTTAGAGGAGAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
Db 1237 ACTTTATGGAATAATTTCACTGGAAGGAAAGACTAACTTTCTTTGAGAAGAGTAGGCG 1296
Qy 989 AGTATCAGCGTTTTCAGTATGAGTATGCGCAGAAACACAGATAACGTTCTTCACTTGGATGAG 1048
Db 1297 AGTATCAGAGGATGGAGTGAATGCAAGTCCAAACAGAGAATTTCTTTTACCTTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 11
US-09-954-456-1169
; Sequence 1169, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; FILE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456

```

/ CURRENT FILING DATE: 2001-09-18
/ PRIOR APPLICATION NUMBER: US/60/233,617
/ PRIOR FILING DATE: 2000-09-18
/ PRIOR APPLICATION NUMBER: US/60/234,052
/ PRIOR FILING DATE: 2000-09-20
/ PRIOR APPLICATION NUMBER: US/60/234,923
/ PRIOR FILING DATE: 2000-09-25
/ PRIOR APPLICATION NUMBER: US/60/235,134
/ PRIOR FILING DATE: 2000-09-25
/ PRIOR APPLICATION NUMBER: US/60/235,637
/ PRIOR FILING DATE: 2000-09-26
/ PRIOR APPLICATION NUMBER: US/60/235,638
/ PRIOR FILING DATE: 2000-09-26
/ PRIOR APPLICATION NUMBER: US/60/235,711
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: US/60/235,720
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: US/60/235,840
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: US/60/235,863
/ PRIOR FILING DATE: 2000-09-27
/ NUMBER OF SEQ ID NOS: 2276
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 1169
/ LENGTH: 2500
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-09-954-456-1169

```

Query Match	54.9%	Score 578.4	DB 9	Length 2500	
Best Local Similarity	75.0%	Pred. No. 1.4e-156			
Matches 723	Conservative	0	Mismatches 241	Indels 0	Gaps 0
QY	89	ATGAAGAGCCACCTCTAAGAAAGACAGTTCTCGCCGGTTTGTTCATCTTTCCAAATCCAGTACC	148		
DB	397	AGGATGAGCGCTGCTGAGNAAAAACCCCGCGCGCTTTGTCTCACTTCCCATCGAGTACC	456		
QY	149	CTGATATTTTGGAAAAATGTTATAAACAGGCACAGCGCTTCCTTTCTGSACACGAGAAGAGTTCG	208		
DB	457	ATGATATCTGGCAGATGTATAGAAGGCGAGAGGCTTCCTTTTGGACCGCGAGGAGGTG	516		
QY	209	ACTTATCAAAAGGATCTCCCTCACTCGAAACAAGCTTAAAGCAGATGAGAAAGTACTTCACT	268		
DB	517	ACCTCTCCAAGGACATCTCAGCATCTGGAAATCCCTGAAACCCGAGGAGAGATATTTTATAT	576		
QY	269	CTCACATCTTAGCGCTTTTTCGACGCCAGTGAATGTAATGTAATGAAAAATTTGGTGGAGC	328		
DB	577	CCCATGTTCTGGCTTTCTTTTCGACGAAGCGATAGTCATGATAATGAAAACTTGGTGGAGC	636		
QY	329	GCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTCCTCA	388		
DB	637	GAATTTAGCCAAGAAGTTCAGATTACAGAAGCCGCGTGTTCATGGCTTCCAAATTTGCCA	696		
QY	389	TCGAGAAATGTTCACTCAGAGATGTACAGTTTGCTGTAGACACTTACATCAGAGATCCCA	448		
DB	697	TGGAACAACATACATCTGAAATGTATAGTCTTTATTTGACACTTACATTAAGATCCCA	756		
QY	449	AGAAAAAGGAAATTTTATTTAATGCAATTAAGAAACCATGCCCTATGTTAAGAAAAAAGCAG	508		
DB	757	AAGAAAAGGAAATTTCTCTTCAATGCCATTTGAAACGATGCCCTGTGTCAAGAAAGAGGCAG	816		
QY	509	ATTGGGCCCTTGGATCGGATAGCAGATAGAAAACTACTTTTGGGGAAGAGTGGTGGCCT	568		
DB	817	ACTGGGCCCTTGGCCTGGATTTGGGGAACAAGAGGCTACCTATGTGTGAACGTTGTGTAGCCCT	876		
QY	569	TTGTGCTGCTAGAAAGAGTTTTCTTCTCAGGATCTTTTGTGCTGTATATTTCTGCTCAAGA	628		
DB	877	TTGTGCTGCTGGAAGGCATTTTCTTTTCCGGTTCTTTTGGCTCGATATTTCTGGCTCAAGA	936		
QY	629	AGAGAGGTCATTGCCAGGACTCAGCTTTTTCGAATGAATCACTCATCAGCAGAGATGAAGGAC	688		
DB	937	AACAGGACTGATGCTCGCCCTCACATTTTCTAATGAACCTATTATAGCAGAGATGAGGGTT	996		

RESULT 12

```

US-09-954-456-1827
; Sequence 1827, Application US/09954456
; Patent No. US0202115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1827
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1827

```

Query Match 54.9%; Score 578.4; DB 9; Length 2500;
Best Local Similarity 75.0%; Pred. No. 1.4e-156;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;


```
Db 757 AAGAAAGGGAATTTCTCTCAATGCGCATGAAACGATGCGTTGTGTCAGAAAGAGCGAG 816
QY 509 ATTGGGCTTTGGATAGCAGATAGAAAATCTACTTTTGGGGAAGAGTGGGCGCT 568
Db 817 ACTGGGCTTTGGCTGGATGGGGAACAAGAGCTACCTATGTGAACGTGTGTAGCCT 876
QY 569 TTGCTGCTGTAGAAGAGTTTCTCTCAGGATCTTTTGGCTGTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGCACTGGAAGGCAATTTCTTTTCCGGTTCTTTTGGCTGATATTTCTGGCTCAAGA 936
QY 629 AGAGAGCTTTATGCCAGGACTCACTTTTCCAAATGAATCACTACAGCAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCGCTGCGCTCACATTTTCTAAATGAATCTTATAGCAGATGAGGGTT 996
QY 689 TTCACCTGTCACTTTGCTGCTGATGCTTCCAAATCTTACTTAAATAGCCTTCAGAGAAA 748
Db 997 TACACTGTGATTTTGTCTGCTGATGTTCAAAACACCTGTTACAAACCATCGGAGGAGA 1056
QY 749 GGGTCAGGAGATCATTTGTGATGCTGTCAAAATTTGAGCAGGAGTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATATATCAATGCTGTTCGGATAGAACAGAGTTCTCTCAGGCGCT 1116
QY 809 TGGCAGTTGGCTCTATGGAAATGAATTTGCAATTTTGTATGAACAGTACATTTGATG 868
Db 1117 TGCCTGTGAAGCTCATTTGGGATGAATTTGACACTTAATGAAGCAATACATTTGAGTTGTG 1176
QY 869 CTGACAGATTACTTGGNACTTGGATCTCAAGTCTTCAAGGTTTTCAGGAGAAAATCCTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACCTGGGTTTGTAGCAAGGTTTTCAGAGTAGAAGCCCATTTG 1236
QY 929 ATTTATGAAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAACGAGTTTTCAG 988
Db 1237 ACTTTATGAGAAATATTTCACTGGAAGGAAAGACTAACTTCTTTGAGAAGAGTAGGCG 1296
QY 989 AGTATCAGGTTTTCAGGTTATGGCAGAAACACAGATAACGCTTCTTCACTTTGGATGCGAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGATGTCAAGTCCACAGAGAAATCTTTTACCTTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360
```

RESULT 14

```
US-10-733-878-458
; Sequence 458, Application US/10733878
; Publication No. US20040224408A1
; GENERAL INFORMATION:
; APPLICANT: Jean-Philippe Girard
; APPLICANT: Francois Amalric
; APPLICANT: Myriam Roussigne
; APPLICANT: Thomas Clouaire
; TITLE OF INVENTION: THAP PROTEINS AS NUCLEAR RECEPTORS FOR
; TITLE OF INVENTION: CHEMOKINES AND ROLES IN TRANSCRIPTIONAL REGULATION, CELL
; TITLE OF INVENTION: PROLIFERATION AND CELL DIFFERENTIATION
; FILE REFERENCE: BIOBANK.012A
; CURRENT APPLICATION NUMBER: US/10/733,878
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 60/432699
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/485027
; PRIOR FILING DATE: 2003-07-03
; NUMBER OF SEQ ID NOS: 535
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 458
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-733-878-458
```

Query Match

Best Local Similarity

54.9%; Score 578.4; DB 21; Length 2500;

75.0%; Pred. No. 1.4e-156;

RESULT 15

US-10-843-641A-3751

```
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;
QY 89 ATCAAGAGCCACTCCCTAAGAAAGAGTTCTCGCGTTTGTCTATCTTTTCAATTCAGATACC 148
Db 397 AGATAGAGCCGCTGTCTGAGAGAAACCCCGCGCTTGTCTATCTTTCCCATTCAGATACC 456
QY 149 CTGATATTTTGGAAATATGATATAAACAGGCACAGGCTTCTTCTTGACAGCAGAAAGAGTCG 208
Db 457 ATGATATCTGSCAGATGATATAAGAAAGCAGAGGCTTCTTTTGGACCGCGAGAGGTTG 516
QY 209 ACTTATCAAAAGATCTCCCTCACTGGAACAAAGCTTAAAGCAGATGAGAAGTACTTCTATCT 268
Db 517 ACCTCTCCAAGGACATTCAGCACTTGGGAATCCCTGAAACCCGAGAGAGAGATATTTTATAT 576
QY 269 CTCACATCTTAGCCTTTTTCAGCCAGTCAAGTGAATTTGTAATTTGTAATTTGTTGGAGC 328
Db 577 CCCATGTTCTGGCTTCTTTTGCAGCAAGCGATGECATAGTAAATGAAACACTTGGTGGAGC 636
QY 329 GCTTTAGTCAGGAGGTGCGAGTTCCAGAGGCTCGCTGTTTCTATGCGCTTTCAAAATTTCTCA 388
Db 637 GATTTAGCCAGAAAGTTTCAGATTACAGAGCCGCTGTTCTATGCGCTTCCAAATTTGCCA 696
QY 389 TCAGAAATGTTCACTCAGAGATGTACAGTTTGTGTATAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAACATACATTTCTGAAATGTATAGTCTTCTTATTTGACACTTACATAAAGATCCCA 756
QY 449 AGNAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCCCTATGTTAAGNAAAGAGCAG 508
Db 757 AAGAAAGGGAATTTCTCTTCAATGCGCAATGGAACGATGCGCTTGTGTCAAGAAAGAGCAG 816
QY 509 ATTGGGCTTTGCGATGAGATAGAGATAAGAAATCTACTTTTGGGGAAGAGTGTGGCCT 568
Db 817 ACTGGGCTTTGCGCTGATTTGGGGAACAAGAGGCTACCTATGTTGAACGTTGTAGCCT 876
QY 569 TTGCTGCTGTAGAAGAGTTTCTTCTCAGGATCTTTTGTGCTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGCACTGGAAGGCAATTTCTTTTCGCGTTCTTTTGGCTCGATATTTCTGGCTCAAGA 936
QY 629 AGAGAGCTTTATGCCAGGACTCACTTTTTCAAATGAACCTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCGCTGCGCTCACATTTTCTAATGAACCTTATAGCAGAGATGAGGGTT 996
QY 689 TTCACCTGTCACTTTGCTGCTGATGTTCCAAATCTTACTTAAATAGCCTTTCAGAAAGAA 748
Db 997 TACACTGTGATTTTGTCTGCTGATGTTTCAACACCTGGTACAAACCATCGGAGGAGA 1056
QY 749 GGGTCAGGAGATCATTTGTGATGCTGTCAAAATTTGAGCAGGAGTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATATATCAATGCTGTTCGGATAGAACAGAGTTCTCTCAGTGGCCT 1116
QY 809 TGGCAGTTGGCTCTATGGAAATGAATTTGATGAACAGTACATTTGAGTTTGTAG 868
Db 1117 TGCCTGTGAAGCTCATTTGGGATGAATTTGCACTTAATGAAGCAATACATTTGAGTTGTG 1176
QY 869 CTGACAGATTACTTGTGGAACTTGGATCTCAAGGTTTTCAGGAGGAAATCCTTTT 928
Db 1177 CAGACAGACTTATGCTGGAACCTGGGTTTGTAGCAAGGTTTTCAGAGTAGAAGCCCATTTG 1236
QY 929 ATTTATGAAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAACGAGTTTTCAG 988
Db 1237 ACTTTATGAGAAATATTTCACTGGAAGGAAAGACTAACTTCTTTGAGAAGAGTAGGCG 1296
QY 989 AGTATCAGGTTTTCAGGTTATGGCAGAAACACAGATAACGCTTCTTCACTTTGGATGCGAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGATGTCAAGTCCACAGAGAAATCTTTTACCTTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360
```

/ Sequence 3751, Application US/10843641A
/ Publication No. US20050064454A1
/ GENERAL INFORMATION:
/ APPLICANT: Avalon Pharmaceuticals, Inc.
/ TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
/ FILE REFERENCE: 689290-189
/ CURRENT APPLICATION NUMBER: US/10/843,641A
/ CURRENT FILING DATE: 2004-05-12
/ PRIOR APPLICATION NUMBER: US/09/873,367
/ PRIOR FILING DATE: 2001-06-05
/ PRIOR APPLICATION NUMBER: US/09/954,531
/ PRIOR FILING DATE: 2001-09-18
/ PRIOR APPLICATION NUMBER: US/09/954,456
/ PRIOR FILING DATE: 2001-09-25
/ PRIOR APPLICATION NUMBER: US/09/962,436
/ PRIOR FILING DATE: 2001-09-25
/ PRIOR APPLICATION NUMBER: US/09/962,832
/ PRIOR FILING DATE: 2001-09-25
/ PRIOR APPLICATION NUMBER: US/09/964,824
/ PRIOR FILING DATE: 2001-09-27
/ PRIOR APPLICATION NUMBER: US/09/967,768
/ PRIOR FILING DATE: 2001-09-28
/ PRIOR APPLICATION NUMBER: US/09/968,007
/ PRIOR FILING DATE: 2001-10-02
/ PRIOR APPLICATION NUMBER: US/09/969,347
/ PRIOR FILING DATE: 2001-10-02
/ PRIOR APPLICATION NUMBER: US/09/969,708
/ PRIOR FILING DATE: 2001-10-03
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 8447
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 3751
/ LENGTH: 2500
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-843-641A-3751

Query Match 54.9%; Score 578.4; DB 22; Length 2500;
Best Local Similarity 75.0%; Pred. No. 1.4e-156;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;
QY 89 ATGAAGGACCTCTCTAAGAAAGAGTTCTCGCGGTTTGTCTCATCTTCCAAATCCAGTACC 148
DB 397 AGGATGAGCGCTGTCTGAGAGAAACCCCGCGCTTGTCTCATCTTCCCATCGAGTACC 456
QY 149 CTGATATTTGGAATGTATAACAGGCACAGCTTCTCTTCTGGACAGAGAGGTCTG 208
DB 457 ATGATATCTGGCAGATGTATAAGAGGCGAGAGGCTTCTCTTTGGACCGCGAGGAGTTG 516
QY 209 ACTTATCAAGAGATCTCCCTCTACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTCATCT 268
DB 517 ACCTCTCAAGGACATTTCAAGCCTGGGAAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
QY 269 CTCACATCTTAGCTTTTGGACGAGTGTGAATTTGAAATGAAATTTGGTGGAGC 328
DB 577 CCATGTTCTGCTTCTTTGAGCAGAGCATGGCATAGTAATGAAACTTGGTGGAGC 636
QY 329 GCTTTAGTCAGGAGTGCAGGTTCAGAGGCTCGCTGTTTCTATGGCTTTCAAATCTCA 388
DB 637 GATTTAGCCAAGAGTTTCAGATTACAGAGCCGCTGTTTCTATGGCTTCCAAATTTGCCA 696
QY 389 TCGAGATGTTCACTCAGAGATGTACAGTTTGTGTAGTACACATTCATCAGAGATCCCA 448
DB 697 TGGAAAAACATACATCTGAAATGTATAGTCTTCTTATTGACACTTACATAAAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTTTAAATGCAATGAAACCATCCCTATGTTTAAAGAAAAAGCAG 508
DB 757 AAGAAAGGGAATTTCTCTCAATGCAATGAAACCATGCTTGTGTCAAGAAAGAGGCGAG 816
QY 509 ATTGGGCTTGGGATGAGATAGAAAAATCTACTTTTGGGAAAGAGTGGTGGCCT 568
DB 817 ACTGGGCTTGGCTGGATTTGGGCAAGAGGCTACCTATGCTGAACGTTGTAGCCT 876

QY 569 TTGCTGCTGTAGAGGAGTTTTCTCTCAGGATCTTTTGTCTGTATATTTCTGGCTAAAGA 628
DB 877 TTGCTGCTGTAGAGGAGTTTTCTTTCCGGTCTTTTGTGCTGATATTTCTGGCTCAGA 936
QY 629 AGAGAGGTCTTATGCGCAGGACTCATTTCCTCAATGAATCATCAGCAGAGATGAAGGAC 688
DB 937 AACGAGGACTGATGCTGCGCTCACATTTCTAATGAACCTTATTAGCAGAGATGAGGTT 996
QY 689 TTCACCTGTGACTTTGCTTGGCTGATGTTCCAAATCTTAGTAATAAGCCTTCAGAGAAA 748
DB 997 TACACTGTGATTTTGTCTGCTGATGTTCAAAACACCTGGTACACAAACCATCGAGGAGA 1056
QY 749 GGCTCAGGAGGATCATTTGTTGATGCTGTCAAAATTTGAGCAGGAGTTTTTAAACAGAGCCT 808
DB 1057 GAGTAAGAGAAATAATATCAATGCTGTTCCGATAGAACAGGAGTTCTCTCACTGAGGCT 1116
QY 809 TGCAGTTGGCCTCATTTGGAATGAATTTGCAATTTTGATGAACAGTACATTTGATTTGAG 868
DB 1117 TGCCTGTGAAGCTCATTTGGGATGAATTTGCACTCTAATGAAGCAATACATTTGAGTTTGG 1176
QY 869 CTGACAGATTTACTTTGGAACCTTGGATTTCTCAAGGTTTTTTCAGGCGAGAAATCCTTTTG 928
DB 1177 CAGACAGACTTATGCTGGAACCTGGGTTTTAGCAAGGTTTTTTCAGAGTAGAGAACCCATTTG 1236
QY 929 ATTTTATGGAACACATTTCTTTAGAAGGAAAAACAAATTTCTTTGAGAAAACGAGTTTCAG 988
DB 1237 ACTTTATGAGAAATATTTCACTGGAAGGAAAGAACTAACTTCTTTGAGAAAGAGTAGCG 1296
QY 989 AGTATCAGGCTTTTTCAGATTTATGCGCAGAAACCAAGATTAACGTTCTTCCCTTGGATGAG 1048
DB 1297 AGTATCAGAGGATGGGAGTGTGTCAGTCCAAACAGAGAAATTTCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
DB 1357 ACTT 1360

Search completed: October 30, 2005, 06:24:09
Job time : 679.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 30, 2005, 03:23:39 ; Search time 62.25 Seconds
(without alignments)
13084.616 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1846

Sequence: 1 atggggaccggaaaggcc.....tcaccttgatgcagatttt 1053

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10698228/runat_26102005_100607_3977/app_query.fasta_1.2446
-DB=A_Geneseq_16Dec04 -QFMT=fastan -SUFTX=n2p.rag -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -ENDS=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228 @CGN_1_1_140 @runat_26102005_100607_3977 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=0.5 -XGAPOP=6
-XGAPOP=7 -XGAPOP=10 -XGAPOP=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_16Dec04.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	98.6	351	3 AAY84439	Aay84439 Amino aci
2	1821	98.6	351	4 AAB69050	Aab69050 Human rib
3	1817	98.4	351	4 AAB93497	Aab93497 Human pro
4	1454	78.8	389	4 AAB69051	Aab69051 Human R2
5	1454	78.8	389	6 ABU07433	Abu07433 Protein d
6	1454	78.8	389	7 ADJ68753	Adj68753 Human hea
7	1454	78.8	389	8 ADJ66564	Adj66564 Ribonucle
8	1454	78.8	389	8 ADK70435	Adk70435 Respirato
9	1454	78.8	389	8 ADN03789	Adn03789 Antipsori
10	1454	78.8	389	8 ADN04444	Adn04444 Antipsori

11	1454	78.8	389	8	ADO19226	Ado19226 Human PRO
12	1454	78.8	389	8	ADQ09272	Adq09272 Human RRM
13	1454	78.8	389	8	ABM80258	Abm80258 Tumour-as
14	1454	78.8	413	3	AAB43902	Aab43902 Human can
15	1454	78.8	453	4	AAU28017	Aau28017 Human con
16	1454	78.8	453	4	ABG15287	Abg15287 Novel hum
17	1429	77.4	390	2	AAW41765	Aaw41765 Human rib
18	1425.5	77.2	386	8	ADO57309	Ado57309 Kidney de
19	1348	73.0	393	4	ABB64130	Abb64130 Drosophi
20	1242	67.3	430	8	ADN21469	Adn21469 Bacteri
21	1230	66.6	405	6	ABJ26187	Abj26187 Aspergill
22	1214	65.8	381	8	ADN23033	Adn23033 Bacteri
23	1192.5	64.6	381	6	ABJ25587	Abj25587 Aspergill
24	1192	64.6	413	5	ABP73700	Abp73700 Candida a
25	1176.5	63.7	321	8	ADS21332	Ads21332 Bacteri
26	1169.5	63.4	399	6	ABR53196	AbR53196 Protein s
27	1169.5	63.4	399	7	ADK63240	Adk63240 Disease t
28	1169.5	63.4	399	8	ADS43740	Ads43740 Bacteri
29	1161.5	62.9	391	8	ADS44135	Ads44135 Bacteri
30	1161.5	62.9	401	7	ADB70101	Adb70101 C. neofo
31	1132	61.3	352	3	AG05697	Ag05697 Arabidops
32	1130	61.2	329	3	AGS1528	AgS1528 Arabidops
33	1130	61.2	329	3	AG05699	Ag05699 Arabidops
34	1130	61.2	332	3	AGS1527	AgS1527 Arabidops
35	1130	61.2	332	3	AG05698	Ag05698 Arabidops
36	1117	60.5	347	3	AG14546	Ag14546 Arabidops
37	1112.5	60.3	330	3	AG14548	Ag14548 Arabidops
38	1112.5	60.3	333	3	AG14547	Ag14547 Arabidops
39	1099	59.5	314	3	AGS1529	AgS1529 Arabidops
40	1091	59.1	341	3	AG41395	Ag41395 Arabidops
41	1091	59.1	363	3	AG41394	Ag41394 Arabidops
42	1084.5	58.7	330	3	AG41396	Ag41396 Arabidops
43	1072.5	58.1	274	2	AAW26418	Aaw26418 Swinepox
44	1072.5	58.1	274	4	AAB68239	Aab68239 Protein e
45	977.5	53.0	322	2	AAR72736	Aar72736 Plasmodiu

ALIGNMENTS

RESULT 1

AAV84439

ID AAV84439 standard; protein; 351 AA.

AC AAY84439;

XX 25-JUL-2000 (first entry)

XX Amino acid sequence of a human RNA-associated protein.

XX Human; RNA-associated protein; cell proliferation; cancer; inflammation;
immune response; reproductive disorder; actinic keratosis;
atherosclerosis; arteriosclerosis; buritis; cirrhosis; hepatitis;
mixed connective tissue disease; myelofibrosis; primary thrombocythemia;
paroxysmal nocturnal hemoglobinuria; polycythemia vera; psoriasis;
trauma.

XX Homo sapiens.

XX Key Location/Qualifiers

PH Modified-site 18
FT Modified-site /note= "potential phosphorylation site"
FT Modified-site 23
FT Modified-site /note= "potential glycosylation site"
FT Modified-site 29
FT Modified-site /note= "potential phosphorylation site"
FT Modified-site 38
FT Modified-site /note= "potential phosphorylation site"
FT Modified-site 39
FT Modified-site /note= "potential phosphorylation site"
FT Modified-site 65
FT Modified-site /note= "potential phosphorylation site"
FT Modified-site 87
FT Modified-site /note= "potential phosphorylation site"

FT Modified-site 182 /note= "potential phosphorylation site"
 FT Modified-site 225 /note= "potential phosphorylation site"
 FT Modified-site 314 /note= "potential glycosylation site"
 FT Modified-site 329 /note= "potential phosphorylation site"
 XX
 FN WO200015799-A2.
 XX PD 23-MAR-2000.
 XX PF 17-SEP-1999; 99WO-US021688.
 XX PR 17-SEP-1998; 98US-00156039.
 PR 22-SEP-1998; 98US-00158720.
 PR 04-NOV-1998; 98US-00186815.
 PR 08-APR-1999; 99US-0128660P.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Tang YT, Corley NC, Guegler KJ, Gorgone GA, Patterson C;
 PI Hillman JL, Baughn MR, Lai P, Azimzai Y, Yue H, Yang J;
 XX
 DR WPI; 2000-271437/23.
 DR N-PSDB; ARA12411.
 XX
 XX New polypeptides and polynucleotides, useful for preventing and treating
 FT a disorder associated with increased or decreased expression of RNA
 PT associated proteins.
 PT
 XX
 PS Claim 1; Page 103-104; 131pp; English.
 XX
 CC The present sequence represents a human RNA-associated protein. The
 CC expression of RNA-associated proteins is closely associated with
 CC reproductive tissues, nervous tissues, cell proliferation including
 CC cancer, inflammation and immune responses, and so they may be used for
 CC diagnosis, treatment or prevention of cell proliferative,
 CC immune/inflammatory disorders, and reproductive disorders. Diseases and
 CC disorders which may be treated include actinic keratosis,
 CC atherosclerosis, arteriosclerosis, bursitis, cirrhosis, hepatitis, mixed
 CC connective tissue disease, myelofibrosis, paroxysmal nocturnal
 CC hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia
 CC and cancers, and trauma
 XX
 SQ Sequence 351 AA;
 Alignment Scores:
 Pred. No.: 1,03e-209 Length: 351
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.65% Indels: 0
 DB: 3 Gaps: 0
 US-10-698-228-2 (1-1053) x AAY84439 (1-351)
 QY 1 ATGGCGACCCGGAAGCGCGGCTGGATCAGATCAGATCATCTTCA 60
 DB 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
 QY 61 GACACCAACGAAGTGAATAAGTCAATGAAGACCACTCTTAAGAAAGTCTCGC 120
 DB 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
 QY 121 CGGTTTGTCTCTTCCATCCAGTACCTGATATTGCAAAATCTATAAAGCAGCACAG 180
 DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
 QY 181 GCTTCCTCTCGACAGCAGAGAGGTTCGACTTATCAAGCATCTCCCTCACTGGGAACAG 240
 DB 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80

QY 241 CTTAAAGCAGATCAGAGAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCGAGCCAGTGAT 300
 DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
 QY 301 GGAATTTGTAATCAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCAGAGCT 360
 DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 QY 361 CGCTGTTTCTATGGCTTTCAAATCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTTG 420
 DB 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTAATGCAATTGAA 480
 DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCCCTTCGATGATGATGATGATGATAA 540
 DB 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 QY 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTTCCTTCTCAGGA 600
 DB 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 QY 601 TCTTTTCTGCTATATTCTGGCTTAAAGACAGAGGTCTTATGCCAGGACTCACATTTTCC 660
 DB 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCCTCTGTGACTTCTGCTCCCTGATCTCAA 720
 DB 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 QY 721 TACTTAGTAAATAGCCTTCAGAAAGAGGTTCAGGAGATCATTTGTTGCTGTCAAA 780
 DB 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 QY 781 ATTGACACGAGTCTTTTAAACAGAGCCTTCGCCAGTTCCTCATTTGGAATGAATTGCATT 840
 DB 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 QY 841 TTGATGAACACAGTACATTCAGTTTGTAGCTGCAGAGATTACTTGTGAACTTGGATTCTCA 900
 DB 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 QY 901 AAGTTTTTTCAGCAGAAAATCCTTTTGATTTTATGGAACAACTTCTTTAGAAAGGAAAA 960
 DB 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 QY 961 ACAATTTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTCCTGAGTTATGGCAGAAACC 1020
 DB 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 QY 1021 ACAGATAACGCTCTTCACCTTGGATGCAGATTTT 1053
 DB 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 RESULT 2
 ID AAB69050
 XX AAB69050 standard; protein; 351 AA.
 AC AAB69050;
 XX
 DT 18-APR-2001 (first entry)
 DE Human ribonucleotide reductase TP53R2H protein sequence SEQ ID NO:1.
 XX Human; ribonucleotide reductase; cancer; DNA repair; p53.
 OS Homo sapiens.
 XX
 PN WO200100799-A1.
 XX

PD 04-JAN-2001.
 XX
 PF 27-JUN-2000; 2000WO-JP004189.
 PR 28-JUN-1999; 99JP-00181131.
 PR 06-JUL-1999; 99JP-00192391.
 PR 21-JAN-2000; 2000JP-00017770.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 PA (NAKA/) NAKAMURA Y.
 XX
 PI Nakamura Y, Arakawa H, Tanaka H;
 XX N-PSDB; AAF32438.
 DR WPI; 2001-112446/12.
 DR
 XX
 PT Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
 PT diagnosis, treatment and prevention of cancer.
 XX
 PS Claim 1; Fig 1-3; 102pp; Japanese.
 CC The present sequence represents a human ribonucleotide reductase
 CC designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of
 CC the DNA repair mechanism and its activity is induced by p53. It can be
 CC used for the treatment, prevention and diagnosis of a wide range of
 CC cancers
 CC
 XX SQ Sequence 351 AA;
 Alignment Scores:
 Pred. No.: 1.03e-209 Length: 351
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.65% Indels: 0
 DB: 4 Gaps: 0
 US-10-698-228-2 (1-1053) x AAB69050 (1-351)
 QY 1 ATGGGGGACCCCGAAGAGCGCGGCTGGATCAGATGAGATGATCATTTCA 60
 DB 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
 QY 61 GACACCAAGCAAGTCAATTAAGTCAATGAAGACCCACTCTAGAAAGAGTTCTCC 120
 DB 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
 QY 121 CGGTTTGTTCATCTTCCAAATCCAGTACCTCGATATTTGAAAATGTATAAAGCAGCAG 180
 DB 41 ArgPheValIlePheProIleGlnTyrProAspIleIleTyrPheMetTyrLysGlnAlaGln 60
 QY 181 GTTCTCTCTGACAGCAGAGAGTTCGACTTATCAAGAGTCTCCCTCACTGGACAAAG 240
 DB 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 QY 241 CTTAAGCAGATGAGAGTACTTCTCTCATCTTACATCTTACCTTTTCCAGCCAGTAT 300
 DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
 QY 301 GGAATTGTAATGAAAATTTGTGTGAGGCGCTTTAGTCAGAGGTGCGAGTTCAGAGGCT 360
 DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 QY 361 CGCTGTTCTTATGGCTTTCAAAATTCATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 420
 DB 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 QY 421 CTGATAGACACTTACATCAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 480
 DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 QY 481 ACCATGCCCTATGTTAAGAAAAACAGATTCGGCTTCGGATGATAGCAGATAGAAA 540

Db 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 QY 541 TCTACTTTTGGGAAAGAGTGGCTTCTGCTGTAGAGAGAGTCTTCTTCTCAGCA 600
 Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 QY 601 TCTTTTCTGCTATATTTCTGGCTAAAGAAGAGAGTCTTATGCCAGGACTCTCACTTTTCC 660
 Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 QY 661 AATGAATCATCAGCAGATGAGGACTTTCACCTGTGACTTGTGCTGTGCTGTGCTGAA 720
 Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 QY 721 TACTTAGTAATAGCTTCCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGCTCAA 780
 Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 QY 781 ATTGAGCAGGAGTTTTAAACAGAGCCCTTGCAGTTGGCCCTCATTTGGAATGAATGCAAT 840
 Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 QY 841 TTGATGAACAGTACATGAGTTTGTAGTCAAGATTTACTTTGCGAATCTTGATTTCTCA 900
 Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluGlyPheSer 300
 QY 901 AAGTTTTTTCAGGCAAGAAATCCTTTTGTATTTTATGAAACATTTCTTTAGAACGAAA 960
 Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 QY 961 ACAAAATTTCTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTTGCAGTTATGSCAGAAACC 1020
 Db 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 QY 1021 ACAGATAACGCTTTCACCTTGGATGCAAGATTTT 1053
 Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 RESULT 3
 AAB93497
 ID AAB93497 standard; protein; 351 AA.
 XX
 AC AAB93497;
 XX
 DT 26-JUN-2001 (first entry)
 DE Human protein sequence SEQ ID NO:12811.
 XX
 DE Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-00116126.
 XX
 PR 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 XX WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or

PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.

PS Claim 8; SEQ ID NO 12811; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-
XX length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialized methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention

XX Sequence 351 AA;

SQ Alignment Scores:

Pred. No.: 3,13e-209 Length: 351
Score: 1817.00 Matches: 350
Percent Similarity: 100.00% Conservatives: 1
Best Local Similarity: 99.72% Mismatches: 0
Query Match: 98.43% Indels: 0
DB: 4 Gaps: 0

US-10-698-228-2 (1-1053) x AAB93497 (1-351)

QY 1 ATGGCGGACCCGGAAGCGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAAGCAAGTGAATTAAGTCAATGAAGCCACTCTAAGAAAGTCTCCG 120
DB 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGTTTGTTCATCTTCCCAATCCAGTACCCTGATATTGGAAATGTATAAACAGGCACAG 180
DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTTCTGACAGCAGAGAGTTCGACTTATCAAGGATCTCCCTCAGTGGAAACAG 240
DB 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAGCAGATGACAGTACTTCTCTCAGATCTAGCCCTTTTCCAGCCAGTAT 300
DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTTGTAATGAAAATTTGGTGGAGCCCTTTAGTCAGAGGTGCGAGTCCAGAGGCT 360
DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 420
DB 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATGAA 480
DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGGATGGATACAGATAGAAA 540

DB 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTAGAAAGAGTTTCTTCTCAGGA 600
DB 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTGCTATATTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCAGCTTTTCC 660
DB 201 SerPheAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 ATGAACTCATCAGCAGAGATGAAGACTTCACGTGACTTTGCTGCTGCTGCTGCTGCTCAA 720
DB 221 AsnGluIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAAGCTTCAGAAAGAGGCTCAGGAGATCATTTGTTGCTGCTGCTCAA 780
DB 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGACGAGAGTCTTTTAAACAGAGCTTCCAGTTCGCCCTCATTTGGAATGAATTGCAAT 840
DB 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyValAsnCysIle 280
QY 841 TTGATGAAACAGTACATTTAGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGATTTCTCA 900
DB 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGTAAGAAACATTTCTTTAGAAAGGAAA 960
DB 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGATTTATGCCAGAAACC 1020
DB 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGTTCTTCACTTCTGATGAGATTTT 1053
DB 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 4
AAB69051
ID AAB69051 standard; protein; 389 AA.
XX
AC AAB69051;
XX
DT 18-APR-2001 (first entry)
XX
DE Human R2 protein sequence SEQ ID NO:5.
XX
KW Human; ribonucleotide reductase; cancer; DNA repair; p53.
XX
OS Homo sapiens.
XX
PN WO200100799-A1.
XX
PD 04-JAN-2001.
XX
PF 27-JUN-2000; 2000WO-JP004189.
XX
PR 28-JUN-1999; 99JP-00181131.
PR 06-JUL-1999; 99JP-00192391.
PR 21-JAN-2000; 2000JP-00017770.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
PA (NAKA/) NAKAMURA Y.
XX
XX Nakamura Y, Arakawa H, Tanaka H;
XX
DR WPI; 2001-112446/12.
XX
XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
PT diagnosis, treatment and prevention of cancer.
XX


```
QY 91 GAAGAGCCACTCTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTTCCAAATCCAGTACCCT 150
DB 69 AspGluProLeuLeuArgGluAsnProArgArgPheValIlePheProIleGluTyrHis 88
QY 151 GATATTGGAAATGTATAAACAGGACAGGCTTCTCTGAGCAGAGAGAGGTCGAC 210
DB 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
QY 211 TTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATCAGAGTACTTCACTCT 270
DB 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTTAGCCCTTTTTCAGCCAGTGTGAATTTGTAATCAAAATTTGGTGGAGCGC 330
DB 129 HisValLeuAlaPhePheAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY 331 TTATGTCAGAGGTGAGGTTCAGAGGCTCGCTTCTATGGCTTTCAAATTTCTCATC 390
DB 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCAAG 450
DB 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAAGGGAATTTTATTAATGCAATTAAGAACCATGCCCTATCTTTAAAGAAAGAGAGAT 510
DB 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
QY 511 TGGGCTTCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 570
DB 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 228
QY 571 GCTGCTGATAGAGAGTTTCTTCAGGATCTTTGCTGCTATATCTGCTTAAGAGAG 630
DB 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGGTCTTATGCCAGGACTCTTTTCCATGATGATGATGATGATGATGATGATGATGAT 690
DB 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY 691 CACTGTGACTTTGCTTGCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 750
DB 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGAGATCATTTGATGCTGCTCAAAATGAGCAGAGTTTAAACAGAGCCTTG 810
DB 289 ValArgGluIleIleAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCCTCATTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 870
DB 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
QY 871 GACAGATTACTTGTGGAATCTGATTTCTCAAAAGTTTTCAGGACAGAAATCTTTTGTAT 930
DB 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY 931 TTATGGAACAACTTTCTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAGAG 990
DB 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY 991 TATCAGCCTTTTGGAGTTATGTCAGAAACCAACAGATTAAGCTTTCACCTTGGATGAGAT 1050
DB 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY 1051 TTT 1053
DB 389 Phe 389
RESULT 9
ADN03789
ID ADN03789 standard; protein; 389 AA.
XX
```

```
AC ADN03789;
XX 01-JUL-2004 (first entry)
XX Antipsoriatic protein sequence #91.
XX antipsoriatic; gene therapy; psoriasis; diagnosis.
XX Homo sapiens.
XX WO2004028479-A2.
XX 08-APR-2004.
XX 25-SEP-2003; 2003WO-US030907.
XX 25-SEP-2002; 2002US-0414006P.
XX (GETH ) GENENTECH INC.
XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
XX Wu TD;
XX WPI; 2004-305105/28.
XX N-PSDB; ADN03788.
XX New PRO nucleic acid or polypeptide, useful for preparing a
XX pharmaceutical composition for diagnosing or treating psoriasis in a
XX mammal.
XX Claim 9; SEQ ID NO 183; 3069pp; English.
XX The invention relates to novel polynucleotide and polypeptides for
XX treating psoriasis or a sequence having at least 80% identity to the
XX above sequences. The nucleic acid is useful for preparing a composition
XX for diagnosing or treating psoriasis in a mammal. This sequence
XX corresponds to one of the polypeptides of the invention.
XX Sequence 389 AA;
SQ
Alignment Scores:
Pred. No.: 1,98e-165 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservatives: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 8 Gaps: 0
US-10-698-228-2 (1-1053) x ADN03789 (1-389)
QY 91 GAAGAGCCACTCTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTTCCAAATCCAGTACCCT 150
DB 69 AspGluProLeuLeuArgGluAsnProArgArgPheValIlePheProIleGluTyrHis 88
QY 151 GATATTGGAAATGTATAAACAGGACAGGCTTCTCTGAGCAGAGAGAGGTCGAC 210
DB 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
QY 211 TTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATCAGAGTACTTCACTCTCT 270
DB 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTTAGCCCTTTTTCAGCCAGTGTGAATTTGTAATCAAAATTTGGTGGAGCGC 330
DB 129 HisValLeuAlaPhePheAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY 331 TTATGTCAGAGGTGAGGTTCAGAGGCTCGCTTCTATGGCTTTCAAATTTCTCATC 390
DB 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCAAG 450
DB 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
```


QY 931 TTTATGAAACATTTCTTTAGAGGAAACAAATTTCTTTCAGAAACGAGTTTCAGAG 990
 Db 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPheGluLysArgValGlyGlu 368
 QY 991 TATCAGCGTTTTCAGTTATGCGAGAAACACAGATAACGTTTACCTGGATGAGAT 1050
 Db 369 TyrGlnArgMetGlyValMetSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
 QY 1051 TTT 1053
 Db 389 Phe 389
 RESULT 12
 ID ADQ09272 standard; protein; 389 AA.
 XX ADQ09272;
 AC ADQ09272;
 XX 23-SEP-2004 (first entry)
 XX Human RRM2 protein SEQ ID NO:457.
 DE
 XX thanatos-associated protein; THAP; THAP responsive gene; THAP family;
 KW THAP responsive element; angiogenesis; inflammation; metastasis; cancer;
 KW apoptosis; cardiovascular disease; neurodegenerative disease; chemokine;
 KW antiangiogenic; antiinflammatory; cardiovascular; cytostatic;
 KW neuroprotective; osteopathic; THAP modulator; THAP synthesis modulator;
 KW human.
 XX
 OS Homo sapiens.
 XX
 PN WO2004055050-A2.
 XX
 PD 01-JUL-2004.
 XX
 XX 10-DEC-2003; 2003WO-IB006434.
 PF
 PR 10-DEC-2002; 2002US-0432699P.
 PR 03-JUL-2003; 2003US-0485027P.
 XX
 PA (ENDO-) ENDOCUBE SAS.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 PI Girard J, Amalric P, Roussigne M, Clouaire T;
 DR WPI; 2004-525034/50.
 DR N-PSDB; ADQ09273.
 XX
 PT Modulating expression of a Thanatos (death)-Associated Protein (THAP)
 PT responsive gene for preventing or treating e.g. cancer or inflammation,
 PT comprises modulating the interaction of a THAP polypeptide with a nucleic
 PT acid.
 XX
 PS Example 47; SEQ ID NO 457; 612pp; English.
 XX
 CC The present invention describes a method for modulating the expression of
 CC a thanatos (death)-associated protein (THAP) responsive gene. The method
 CC comprises modulating the interaction of a THAP-family polypeptide or its
 CC biological fragment with a nucleic acid, and so enhancing or repressing
 CC the expression of the THAP responsive gene. Also described: (1) a method
 CC of modulating the expression of a gene responsive to a THAP/chemokine
 CC complex; (2) a pharmaceutical composition comprising a THAP responsive
 CC element in a pharmaceutical carrier; (3) a transcription factor decoy
 CC consisting essentially of a THAP responsive element; (4) a cell
 CC comprising a transcription factor decoy described above; (5) methods of
 CC modulating the interaction between a nucleic acid and a THAP-family
 CC polypeptide or its biological fragment, or a nucleic acid and a
 CC THAP/chemokine complex; (6) a vector packaging cell line comprising a
 CC cell comprising a viral vector which comprises a promoter operably linked
 CC to a nucleic acid encoding a THAP-family polypeptide or its biological
 CC fragment; (7) a method of constructing a cell which expresses a
 CC recombinant THAP-family polypeptide; (8) a method of ameliorating
 CC symptoms associated with a condition mediated by a THAP/chemokine complex

CC ; (9) methods of identifying a test compound that modulates transcription
 CC at a THAP responsive element or that modulates the transport of a
 CC chemokine into the nucleus; (10) methods for reducing the symptoms
 CC associated with a condition selected from excessive or insufficient
 CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
 CC or insufficient apoptosis, cardiovascular disease and neurodegenerative
 CC diseases; symptoms associated with a condition resulting from the
 CC activity of a chemokine or a THAP-family polypeptide in an individual; or
 CC symptoms associated with transcriptional repression or activation
 CC mediated by a THAP-family polypeptide in an individual; (11) a vector
 CC comprising a THAP responsive promoter operably linked to a nucleic acid
 CC encoding a detectable product; (12) a genetically engineered cell
 CC comprising the vector described above or that expresses a THAP-family
 CC polypeptide or its biological fragment; (13) an in vitro transcription
 CC reaction comprising a nucleic acid comprising a THAP responsive promoter,
 CC ribonucleotides and an RNA polymerase; and (14) an isolated mutant THAP-
 CC family polypeptide that does not bind to a chemokine. The pharmaceutical
 CC composition has antiangiogenic, antiinflammatory, cardiovascular,
 CC cytostatic, neuroprotective and osteopathic activities, and can be used
 CC as a THAP and THAP synthesis modulator. The composition can be used for
 CC modulating the expression of a THAP responsive gene. Modulation is useful
 CC for reducing symptoms of conditions such as excessive or insufficient
 CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
 CC or insufficient apoptosis, cardiovascular disease or neurodegenerative
 CC diseases. The present sequence is used in the exemplification of the
 CC present invention.
 XX
 SQ Sequence 389 AA;

Alignment Scores:

Pred. No.: 1,98e-165 Length: 389
 Score: 1454.00 Matches: 269
 Percent Similarity: 94.08% Conservative: 33
 Best Local Similarity: 83.80% Mismatches: 19
 Query Match: 78.76% Indels: 0
 DB: 8 Gaps: 0

US-10-698-228-2 (1-1053) x ADQ09272 (1-389)

QY 91 GAAGAGCCACTCTCTAAGAAAGAGTTCTCGCGGTTGTCTATTTTCCAAATCCAGTACCCT 150
 Db 69 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 88
 QY 151 GATATTTGGAATGTATTAACAGGACAGGCTTCTCTCGGACAGCAGAGAGGTGCAC 210
 Db 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
 QY 211 TTATCAAGGATCTCCTCACTGGAACAGCTTAAAGCAGATGAGAGTACTTCTCATCTCT 270
 Db 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
 QY 271 CACATCTTAGCCCTTTTTCAGCCAGTGATGGAATTTGAAATTTGAAATTTGGTGAGGCGC 330
 Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
 QY 331 TTATGTCAGGAGTGAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAATTTCTCATC 390
 Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
 QY 391 GAGAAATGTTCCACTCAGAGATGTACAGTTTCTGATGACACTTACATCAGAGATCCCAAG 450
 Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
 QY 451 AAAAGGGAATTTTATTTAATGCAATTTGAACACCATGCCCTATGTATAAGAAAAAGCAGAT 510
 Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
 QY 511 TGGGCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 570
 Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValValAlaPhe 228
 QY 571 GCTGCTGTAGAAGGATTTTCTCTCAGGATCTTTTCTGCTGCTATATTCTGCTAAAGAG 630

229	AlaAlaValGluGlyIlePhePheSerClySerPheAlaSerIlePheTrpLeuLysLys	248
631	AGAGGTCTTATGCCAGGACTCACCTTTTCCAAATGAATCATCAGCAGAGATGAAGACTT	690
249	ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu	268
691	CACGTGTGACTTTGCTTGCCTGATGTTCCAAATCTTGTAAATATAGCCTTCAGAAAGAGG	750
269	HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg	288
751	GTCCAGGAGATCATTTGTTGATGCTGTCCAAATTCAGCAGGAGTGTAAACAGAGACCTTG	810
289	ValArgGluIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu	308
811	CCAGTGGCCCTCATTTGGAATCAATTCGATTTTGATGAACAGTACATTCAGTTGTAGCT	870
309	ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla	328
871	GACAGATTACTTGTGGAACTCTGGATTTCTCAAAGGTTTTTCAGGCAGAAAAATCCCTTTTGAT	930
329	AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp	348
931	TTTATGAAAAACATTTCTTTAGAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAGAG	990
349	PheMetGluAsnIleSerLeuGluGlyLysThrAsnPheGluLysArgValGlyGlu	368
991	TATCAGCGTTTTGCAGTTATGGCAGAAACACAGATAACGTCCTTCCTTCGGATCGAGAT	1050
369	TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp	388
1051	TTT 1053	
389	Phe 389	

RESULT 13
ABM80258

ID	ABM80258	standard; protein; 389 AA.
XX		
XX	ABM80258;	
XX		
AC		
XX		
DT	18-NOV-2004	(first entry)
XX		
XX		
DE	Tumour-associated antigenic target (TAT)	polypeptide PR062893, SEQ:652.
XX		
XX	Tumour-associated antigenic target; TAT; human; overexpression; cancer;	
KW	tumour; diagnosis; cell proliferative disorder; breast cancer;	
KW	colorectal cancer; lung cancer; ovarian cancer; liver cancer;	
KW	central nervous system cancer; bladder cancer; pancreatic cancer;	
KW	cervical cancer; melanoma; leukaemia; hybridisation probe;	
KW	chromosome identification; chromosome mapping; gene mapping;	
KW	gene therapy; cytostatic.	
XX		
XX		
OS	Homo sapiens.	
XX		
XX	WO2004030615-A2.	
XX		
XX	15-APR-2004.	
PD		
XX		
XX	29-SEP-2003; 2003WO-US028547.	
PF		
XX		
XX	02-OCT-2002; 2002US-0414971P.	
PR		
XX	(GETH) GENENTECH INC.	
XX		
PA		
XX		
XX	Wu TD, Zhang Z, Zhou Y;	
PI		
XX	WPI; 2004-347921/32.	
DR		
DR	N-PSDB; ACN37637.	
XX		
XX		
PT	New tumor-associated antigenic target polypeptides and nucleic acids,	
PT	useful in preparing a medicament for treating or detecting a	
PT	proliferative disorder, e.g. breast, lung, colorectal, ovarian or	
PT	prostate cancer or tumor.	

Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
Qy 691 CACTGTGACTTTGCTTGCTGATGTCCAACTACTAGTAATAAGCCTTCAGAGAAAG 750
Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
Qy 751 GTCAGGAGATCAATGTTGATGCTGTCAAAATTTGAGCAGAGATTTTAAACAGAGCCTTG 810
Db 289 ValArgGluIleIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
Qy 811 CCAGTTGGCCTCATTTGGAATGAATTCATTTTGATGAAACAGTACATTGAGTTTGTAGCT 870
Db 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
Qy 871 GACAGATTACTTGTGGAACTTGATCTCAAGGTTTTCAGGCAGAAATCTTTTGAT 930
Db 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
Qy 931 TTTATGGAAAACATTTCTTTAGAGGAAAACAAATTTCTTTGAGAAACGAGTTTCAGAG 990
Db 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
Qy 991 TATCAGCGTTTTCAGTTTATGGCAGAAACACAGATAACGTTCTTCACCTTGGATGCAGAT 1050
Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
Qy 1051 TTT 1053
Db 389 Phe 389
RESULT 14
AAB43902
ID AAB43902 standard; protein; 413 AA.
XX
AC AAB43902;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human cancer associated protein sequence SEQ ID NO:1347.
XX
KW Human; cancer associated gene; cancer antigen; detection; cancer;
KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
KW dermatological; neuroprotective; thrombolytic; coagulant; neutropic;
KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW allergic reaction; graft versus host disease; organ rejection;
KW haemostatic; thrombolytic; cardiovascular disorder; infection;
KW neurological disease; drug screening.
XX
OS Homo sapiens.
XX
PN WO200055350-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US005882.
XX
PR 12-MAR-1999; 99US-0124270P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-587533/55.
DR N-PSDB; AAC78111.
XX
PT Novel isolated nucleic acids comprising sequences encoding peptides
XX useful for treating or diagnosing e.g. cancer.
XX
PS Claim 11; Page 2000-2002; 2352pp; English.

XX AAC77607 to AAC78448 encode the human cancer associated proteins given in
CC AAB43398 to AAB44239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vulnery; immunomodulator;
CC antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC dermatological; neuroprotective; cardiac; thrombolytic; coagulant;
CC neutropic; vasotropic; antipsoriatic and antiangiogenic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention
XX
SQ Sequence 413 AA;
Alignment Scores:
Pred. No.: 2,04e-165 Length: 413
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 3 Gaps: 0
US-10-698-228-2 (1-1053) x AAB43902 (1-413)
Qy 91 GAAGAGCCACTCTCTAAGAAAGAGTTCTCGCGGTTTGTCTCATTTTCCAATCCAGTACCCT 150
Db 93 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 112
Qy 151 GATATTGGAAATGTATAAACAGGCACAGGCTTCCTCTGGACAGACAGAGAGTCCAC 210
Db 113 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 132
Qy 211 TTATCAAGAGTCTCCCTCACTCGAACAGCTTAAAGCAGATGAGAGTACTTCATCTCT 270
Db 133 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgGlyPheIleSer 152
Qy 271 CACATCTTAGCCCTTTTTCAGCCAGTGTGATGGAATTTGAAATGAAATTTGGTGGAGCGC 330
Db 153 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 172
Qy 331 TTTAGTCAGAGGTGAGGTTCCAGAGCTCGCTGTTTCTATGCTTTCAAATTCATC 390
Db 173 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 192
Qy 391 GAGAAATGTTCACTCAGAGATGTACAGTTTCTGTGATGACACTTACATCAGAGATCCCAAG 450
Db 193 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 212
Qy 451 AAAAGGGAATTTTATTATTAATGCAATTAACCAACCATGCCCTATGTAAGAAAAAGCAGAT 510
Db 213 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 232
Qy 511 TGGGCTTCGATGGATGAGATAGAGAAATCTACTTTTGGGAAAGAGTGGTGGCCTTT 570
Db 233 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaAlaPhe 252
Qy 571 GCTGCTGTAGAAGAGTTTCTTCTCAGGATCTTTTGTCTGCTATATTCTGCTAAAGAG 630
Db 253 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 272
Qy 631 AGAGGTCTTATGCCAGGACTCACCTTTTCCAACTCACTCAGCAGAGATGAAGACTT 690

Db 273 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 292
QY 691 CACTGTGACTTTCCTTCCTGATGTTCCCAATCTAGTAATAAGCCTTCAGAGAAGG 750
Db 293 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 312
QY 751 GTCAGGAGATCATTTGTCGATGCTGCAAAATGAGCAGAGATTTTAAACAGAAGCCTTG 810
Db 313 ValArgGluIleIleAsnAlaValArgIleGluGluPheLeuThrGluAlaLeu 332
QY 811 CCAGTGGCTCATGGAATGATTCATTTGATCAAAACAGTACATTCAGTTGTAGCT 870
Db 333 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnIleGluPheValAla 352
QY 871 GACAGATTACTTGTGGAATCTGATCTCAAGGTTTTCAGCAGAAATCTTTTGAT 930
Db 353 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 372
QY 931 TTTATGGAAAAATTTCTTTAGAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAGAG 990
Db 373 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 392
QY 991 TATCAGCTTTGTCAGTTATGTCAGAAACACAGATAAGCTTTCACCTTGATGAGAT 1050
Db 393 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 412
QY 1051 TTT 1053
Db 413 Phe 413
RESULT 15
AAU28017
ID AAU28017 standard; protein; 453 AA.
AC AAU28017;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human contig polypeptide sequence #170.
XX
KW Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
KW nervous system disorder; inflammatory disorder; cell differentiation;
KW angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;
KW cytostatic; antirheumatic; antiarthritic; vulnery; antiinflammatory;
KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
KW neuroprotective; osteopathic; antidiabetic; antiallergic;
KW immunostimulant; analgesic; gene therapy.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200164834-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US004926.
XX
PR 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
PR 17-JUN-2000; 2000US-00597707.
PR 14-JUL-2000; 2000US-00616807.
PR 19-SEP-2000; 2000US-00664641.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao Q, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;
PI Drmanac R;
XX
XX WPI; 2001-589862/66.

DR N-PSDB; AAS44917.
XX Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis, treatment of cancer,
PT neurological, inflammatory disorders and for use in arrays for detection.
XX
PS Claim 10; Page 146-147; 153pp; English.
XX
CC Sequences AAU27676-AAU28019 represent full-length polypeptides and contig
CC polypeptides of the invention. The proteins and their associated DNA
CC sequences are useful for the treatment, diagnosis and prevention of
CC various types of disorder in a mammalian subject such as a human, dog,
CC monkey, mouse, hamster or rat. The disorders include cancers such as
CC leukaemia, lymphoma and neuroblastoma, autoimmune disorders such as
CC multiple sclerosis, connective tissue disease, rheumatoid arthritis,
CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's
CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
CC bowel disease. The sequences exhibit activity relating to angiogenesis,
CC cell proliferation, cell differentiation, stem cell growth factor,
CC activin or inhibin. Therefore, they can be used to manipulate stem cells
CC in culture to give rise to neuroepithelial cells that can be used to
CC augment or replace cells damaged by illness, accidental damage or genetic
CC disorders. The sequences may also be used for regeneration of bone,
CC cartilage, tendons and ligaments and in tissue repair and burn healing.
CC Note: Some sequences for this patent did not form part of the printed
CC specification, but were obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 453 AA;
Alignment Scores:
Pred. No.: 2,13e-165 Length: 453
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: Gaps: 0
US-10-698-228-2 (1-1053) x AAU28017 (1-453)
QY 91 GAGAGCCACTCTTAAAGAAAGATCTTCGCGGTTTGTTCATCTTCCATCAGTACCCT 150
Db 133 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 152
QY 151 GATATTGGAAAAATGTATAACACAGGCACAGCTTCTCTCGACAGCAGAGAGTCCGAC 210
Db 153 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheThrAlaGluGluValArg 172
QY 211 TTATCAAGGATCTCCCTCCTCAGTGGAAACAGCTTAAAGCAGATGAGAGTACTCTCTCT 270
Db 173 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 192
QY 271 CACATCTTAGCCTTTTTCAGCCAGTATGGAATGTAAATGAAATTTGGTGGAGCGC 330
Db 193 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 212
QY 331 TTTTAGTCAGGAGTGCAGGTTCCAGAGCGCTCGCTGTTTCTATGGCTTTCAAAATCTCATC 390
Db 213 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 232
QY 391 GAGAATGTTTCACCTCAGAGATGTACAGTTTGTGTATAGACACTTACATCAGAGATCCCAAG 450
Db 233 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 252
QY 451 AAAGGGAATTTTATTATTAATGAAATTCACCATTCCTTATGTTTAAAGAAAAACAGAT 510
Db 253 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 272
QY 511 TGGGCTTCGATGGATAGCAGATAGAAATCTACTTTTGGGAAAGAGTGTGTCCTTT 570

Db 273 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrglyGluArgValValAlaPhe 292
QY |||||GCTGCTGTAGAGGAGTTCTTCTCAGGATCTTTTGGCTGTATATCTGGCTTAAAGAG 630
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 293 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 312
QY 631 AGAGTCTTATGCCAGGACTCACTTTTCCAAATGAATCATCAGCAGAGATGAAGGACTT 690
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 313 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 332
QY 691 CACTGTGACTTTGCTTGCTGTGATGTTCCAAATCTTAGTAATAAGCCTTCAGAGAAAGG 750
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 333 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 352
QY 751 GTCAGGGAGATCATGTTGATGCTGTCAAAATTGACGAGGAGTCTTTTAAACAGAGCCTTG 810
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 353 ValArgGluIleIleGlnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 372
QY 811 CCAGTTGGCCTCATTTGGAATGAATTCATTTTGTATGAAACAGTACATTCAGTTTGTAGCT 870
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 373 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrlleGluPheValAla 392
QY 871 GACAGATTACTTGTGGAACTTGGATCTCAAGGTTTTCAGGCGAGAAATCCTTTTGAT 930
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 393 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 412
QY 931 TTTATGGAACACATTTCTTAGAGGMAAACAAATTTCTTTGAGAAACGAGTTTCAGAG 990
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 413 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 432
QY 991 TATCAGCGTTTTCAGTATGTCAGAAACACAGATTAACGTCTTCACCTTGGATGCGAT 1050
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 433 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 452
QY 1051 TTT 1053
Db ||| 453 Phe 453

Search completed: October 30, 2005, 06:28:41
Job time : 79.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 30, 2005, 05:15:40 ; Search time 15.75 Seconds
(without alignments)

12865.556 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1846

Sequence: 1 atggggccggaaagcc.....tcaccttgatgcagatttt 1053

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10698228/runat_26102005_100608_3995/app_query.fasta_1.2446
-DB=PIR_79 -QFMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228@cgn_1_1_39@runat_26102005_100608_3995 -NCPU=6 -ICPU=3
-NO WMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1821	98.6	366	2 T46249	hypothetical prote
2	1454	78.8	389	2 S25854	ribonucleoside-dip
3	1443	78.2	390	1 S86735	ribonucleoside-dip
4	1421	77.0	386	2 S27153	ribonucleoside-dip
5	1331	72.1	384	2 S24585	ribonucleoside-dip
6	1280	69.3	319	2 T30782	ribonucleoside-dip
7	1280	69.3	319	2 B72154	E4L protein - vari
8	1279	69.3	319	1 RDVZV	ribonucleoside-dip
9	1279	69.3	333	2 T28466	ribonucleoside-dip
10	1266	68.6	333	2 H36839	C8L protein - vari
11	1258.5	68.2	348	2 T30470	ribonucleoside-dip
12	1217	65.9	299	1 ROSS2R	ribonucleoside-dip
13	1214	65.8	381	2 T18876	hypothetical prote
14	1169.5	63.4	399	1 A26916	ribonucleoside-dip

15	1161.5	62.9	391	2 T39992	ribonucleoside-dip
16	1161.5	62.9	391	2 S34808	ribonucleoside-dip
17	1128.5	61.1	329	2 T03688	ribonucleoside-dip
18	1070.5	58.0	340	2 S68538	ribonucleoside-dip
19	1024.5	55.5	349	2 E49412	ribonucleoside-dip
20	950	51.5	324	2 T19778	probable ribonucle
21	805	43.6	345	2 S59744	ribonucleoside-dip
22	681	36.9	668	2 T29884	hypothetical prote
23	638	34.6	327	1 RDVZAS	ribonucleoside-dip
24	385	20.9	331	2 A84389	ribonucleoside red
25	384.5	20.8	415	2 A83502	ribonucleoside red
26	358.5	19.4	346	2 G81728	ribonucleoside-dip
27	357.5	19.4	346	2 E71456	probable ribonucle
28	335	18.1	346	2 B72010	ribonucleoside-dip
29	335	18.1	346	2 F86613	ribonucleoside red
30	329	17.8	306	1 WMBE18	ribonucleoside-dip
31	324	17.6	303	2 T47526	ribonucleoside-dip
32	322	17.4	320	2 T42563	ribonucleoside-dip
33	318	17.2	321	1 WMBE11	ribonucleoside-dip
34	302.5	16.4	314	1 WMBE14	ribonucleoside-dip
35	301	16.3	302	1 WMBE12	ribonucleoside-dip
36	294.5	16.0	347	2 C87281	hypothetical prote
37	292.5	15.8	340	1 WMBE17	ribonucleoside-dip
38	290	15.7	305	2 S55655	ribonucleoside-dip
39	290	15.7	328	2 C97781	hypothetical prote
40	289.5	15.7	337	1 WMBE32	ribonucleoside-dip
41	288.5	15.6	337	1 WMBE12	ribonucleoside-dip
42	288.5	15.6	341	2 C71858	ribonucleoside-dip
43	286.5	15.5	341	2 D64565	ribonucleoside-dip
44	285	15.4	305	2 T03155	ribonucleoside-dip
45	285	15.4	324	2 B71655	ribonucleoside-dip

ALIGNMENTS

RESULT 1

T46249

hypothetical protein DKFZp761E1312.1 - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004

C;Accession: T46249

R;Ansonge, W.; Winkler, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, January 2000

A;Reference number: Z23028

A;Accession: T46249

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-366 <AAA>

A;Cross-references: UNIPROT:Q9NTD8; EMBL:AL137348

A;Experimental source: adult amygdala; clone DKFZp761E1312

C;Genetics:

A;Note: DKFZp761E1312.1

C;Superfamily: ribonucleoside reductase small subunit

Alignment Scores:

Pred. No.:	2.85e-157	Length:	366
Score:	1821.00	Matches:	351
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	98.65%	Indels:	0
DB:	2	Gaps:	0

US-10-698-228-2 (1-1053) x T46249 (1-366)

QY 1 ATGGGCGACCCGAAAGCGCGCGCTGGATCAGGATGAGATCATCTTCA 60

Db 16 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 35

QY 61 GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120

Db 36 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 55

QY 121 CGTTTGTTCATCTTTCATCCAGTACCTCCATGATATTGTGAAATGATATAACGCACAG 180

Db 56 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 75
QY 181 GCTTCTCTTCGACAGCAGAGAGCTCGACTTATCAAGAGATCTCCCTCACTGGACAG 240
Db 76 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 95
QY 241 CTTAAAGCAGATGAGAAGTACTCTCTCATCTCTCATCTCTAGCCCTTTTTCAGCCAGTGT 300
Db 96 LeuLysAlaAspGluLysTyrPheIleSerHisLeuAlaPhePheAlaAlaSerAsp 115
QY 301 GGAATTGTAATGAAATTTGTGTGAGCGCTTTAGTCAGAGGTCAGGTTCCAGAGGT 360
Db 116 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 135
QY 361 CGCTGTTCTATGCTTCTCAATCTCTCATCGAGATGTTCTCATCGAGATGACAGTTTG 420
Db 136 ArgCysPheTyrGlyPheGlnIleuIleGluAsnValHisSerGluMetTyrSerLeu 155
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 156 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 175
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTTCGATGGATGAGATAGAAAA 540
Db 176 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 195
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTCGCTGTAGAGAGGATTTCTTCTCAGA 600
Db 196 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGluValPhePheSerGly 215
QY 601 TCTTTTGTCTATATCTCGGTAAAGAGAGAGGTCTTATCCAGGACTCACTTTTTC 660
Db 216 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 235
QY 661 AATGAACCTCATCAGCAGATGAAGGACTTCTACTGTGACTTTTCTGCTGCTGATTTCCAA 720
Db 236 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 255
QY 721 TACTTAGTAAATAGCCTTCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db 256 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAlaValLys 275
QY 781 ATTGACGAGGAGTTTATACAGAGCCTTGCAGTTGGCTCATTTGGAATGATTCATT 840
Db 276 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 295
QY 841 TTGATGAAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 900
Db 296 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 315
QY 901 AAGGTTTTTCAGGAGAAAATCCTTTTGTATTTATGGAATAACATTCTTTAGAGAAAA 960
Db 316 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 335
QY 961 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGGTTTTCAGGAAACC 1020
Db 336 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 355
QY 1021 ACAGATAACGCTCTTCACTTGGATGATGAGATTTT 1053
Db 356 ThrAspAsnValPheThrLeuAspAlaAspPhe 366

RESULT 2
S25854
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - human
N:Alternate names: ribonucleotide reductase M2 chain; ribonucleotide reductase small chain
C:Species: Homo sapiens (man)
C:Date: 08-Jun-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S25854
R:Pavloff, N.; Rivard, D.; Masson, S.; Shen, S.H.; Mes-Masson, A.M.
DNA Seq. 2, 227-234, 1992
A:Title: Sequence analysis of the large and small subunits of human ribonucleotide reductase

A:Reference number: S25853; MUID:92329977; PMID:1627826
A:Accession: S25854
A:Molecule type: mRNA
A:Residues: 1-389 <PAV>
A:Cross-references: UNIPROT:P31350; EMBL:X59618; MID:g36154; PIDN:CAA42181.1; PID:g36155
C:Genetics:
A:Gene: GDB:RRM2
A:Cross-references: GDB:120358; OMIM:180390
A:Map position: 2p25-2p24
C:Superfamily: ribonucleoside reductase small subunit
C:Keywords: deoxyribonucleotide biosynthesis; iron; metalloprotein; oxidoreductase
F:138,169,172,232,266,269Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, His) #status predicted
F:176/Active site: Tyr (stable tyrosyl radical) #status predicted

Alignment Scores:
Pred. No.: 6,58e-124 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 2 Gaps: 0

US-10-698-228-2 (1-1053) x S25854 (1-389)

QY 91 GAAGAGCCACTCTTAAGAAAGAGTCTTCGCGGTTTGTCTCATCTTCCAATCCAGTACCCT 150
Db 69 AspGluProLeuLeuArgGluAsnProArgArgPheValIlePheProIleGluTyrHis 88
QY 151 GATATTGGAAATGATATAACAGGACAGGCTTCTTCTGGACAGAGAGAGGTTCAC 210
Db 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
QY 211 TTATCAAGGATCTCCCTCATCTGGAACAGCTTAAACAGATGAGAGTACTTCTCATCT 270
Db 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTTAGCCCTTTTTCAGCCAGTCAGTGAATTTAAATGAAATTTGGTGGAGCGC 330
Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyLeuValAsnGluAsnLeuValGluArg 148
QY 331 TTTAGTCAGAGGTGCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTTCAAAATCTCATC 390
Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGAAATGTTCACTCAGAGATGTACAGTTTCTCTATAGACACTTACATCAGAGATCCCAAG 450
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAAGGGAATTTTATTATGCAATTTGAAACCATGCCCCTATGTTAAGAAAAAGCAGAT 510
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
QY 511 TGGCCCTTTCGATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 570
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaAlaPhe 228
QY 571 GCTGCTGTAGAAGAGTTTCTTCTCAGGATCTTTTCTGCTGCTATATTTCTGCTTAAGAG 630
Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGGTTCTTATGCCAGGACTCAGCTTTTCCCAATGAACTCATCAGCAGATGAGAGGACTT 690
Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY 691 CACTGTGACTTTCCTTTCCTGATGTTTCCAAATACTTAAATAGCTTTCAGAGAGAAAGG 750
Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGAGATCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 810
Db 289 ValArgGluIleIleIleAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCCCTCATTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 870

```

Db      309 ProValLeuLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
QY      871 GACAGATTACTGTGGAATTCGATTCCTCAAGGTTTTCAGGCAGAAATCTTTTGAT 930
Db      329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY      931 TTTATCGAAAACATTTCTTTAGAGGAAAACAAATTTCTTTGAGAAAACGAGTTTCAGAG 990
Db      349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY      991 TATCAGCGTTTTCAGTTATGCGAGAAACACACAGATAACGCTTTCACCTTGGATGCGAGAT 1050
Db      369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY      1051 TTT 1053
Db      389 Phe 389

RESULT 3
S06735
ribonucleoside-diphosphate reductase (EC 1.17.4.1) chain M2 - mouse
N:Alternate names: ribonucleotide reductase M2 subunit
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S06735; A24835
R:Thelander, M.; Thelander, L.
EMBO J. 8, 2475-2479, 1989
A:Title: Molecular cloning and expression of the functional gene encoding the M2 subunit
A:Reference number: S06735; MUID:90060004; PMID:2684652
A:Accession: S06735
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-390 <THE>
A:Cross-references: UNIPROT:P11157; EMBL:X15666; NID:g50719; PID:CAA33707.1; PID:g50720
R:Thelander, L.; Berg, P.
Mol. Cell. Biol. 6, 3433-3442, 1986
A:Title: Isolation and characterization of expressible cDNA clones encoding the M1 and M2 subunits
A:Reference number: A24835; MUID:87089677; PMID:3025593
A:Accession: A24835
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-390 <THE2>
A:Cross-references: GB:M14223; NID:g200767; PIDN:AAA0062.1; PID:g200768
A:Note: The authors translated the codon GGG for residue 315 as Glu
C:Genetics:
A:Map position: 12
A:Introns: 33/3; 58/3; 107/3; 146/3; 191/2; 223/1; 267/3; 302/3; 340/3
C:Superfamily: ribonucleoside reductase small subunit
C:Keywords: iron; oxidoreductase
F:139,170,173,233,267,270/Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, His) #sb
F:177/Active site: Tyr (stable tyrosyl radical) #status predicted

Alignment Scores:
Pred. No.: 6,58e-123 Length: 390
Score: 1443.00 Matches: 269
Percent Similarity: 90.62% Conservative: 40
Best Local Similarity: 78.89% Mismatches: 26
Query Match: 78.17% Indels: 6
DB: 1 Gaps: 1

US-10-698-228-2 (1-1053) x S06735 (1-390)

QY      49 AGATCATCTTCAGACACCAACGAAAGTGAATAAGTCA-----AAAT 90
Db      50 ArgIlePheGlnAspSerAlaGluLeuGluSerLysAlaProThrAsnProSerValGlu 69
QY      91 GAAGAGCCACTCTCAAGAAAGATTCTCCCGGTTTGTTCATCTTTCCAAATCCAGTACCTT 150
Db      70 AspGluProLeuLeuArgGluAsnProArgPheValValPheProIleGluTyrHis 89
QY      151 GATATTTCGAAATGTATTAACAGGCACAGGCTTCCTTCGACAGCAAGAGTTCGAC 210

```

```

Db      90 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 109
QY      211 TTATCAAGAGATCCCTCCCTCACTCGAACAAGCTTAAGCAGATGAGAACTACTTCATCTCT 270
Db      110 LeuSerLysAspIleGlnHisTrpGluAlaLeuLysProAspGluArgHisPheIleSer 129
QY      271 CACATCTTACCCCTTTTTCAGCCAGTCATGGAATTTGAAATGAAATTTTGGTGGAGCGC 330
Db      130 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 149
QY      331 TTTAGTCAGAGGTGTCAGGTTCCAGAGCTCGCTGTTCTTATGCTTTCAAATCTCATC 390
Db      150 PheSerGlnGluValGlnValThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 169
QY      391 GAGAAATGTTCACTCAGAGATGTACAGTTTGTGTATAGACACTTACATCAGAGATCCCAAG 450
Db      170 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 189
QY      451 AAAAGGGAATTTTATTATTAATGCAATTAACCAATGCTATGTTTAAAGAAAAAGCAGAT 510
Db      190 GluArgGluTyrLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 209
QY      511 TGGGCTTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 570
Db      210 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaAlaPhe 229
QY      571 GCTGCTGTAGAAGAGTTTCTTCTCAGGATCTTTTCTGCTGCTATATTTCTGCTTAAAGAG 630
Db      230 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 249
QY      631 AGAGGTCTTATGTCAGGATCACTTTTCCAAATGAACTCATCAGCAGAGATGAAAGACTT 690
Db      250 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 269
QY      691 CACTGTGACTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 750
Db      270 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProAlaGluGlnArg 289
QY      751 GTCAGGAGATCATTTGTTGATGCTGTCATAATGAGCAGAGATTTTAAACAGAGCCCTTG 810
Db      290 ValArgGluIleIleThrAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 309
QY      811 CCAGTTGGCTCATTCGAAATGAAATGCTTTCATGATGAAACAGTACATTGAGTTGTAGCT 870
Db      310 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 329
QY      871 GACAGATTACTTGTGGAATTCGATTCCTCAAGGTTTTCAGGCAGAAAATCTTTTGAT 930
Db      330 AspArgLeuMetLeuGluLeuGlyPheAsnLysIlePheArgValGluAsnProPheAsp 349
QY      931 TTTATCGAAAACATTTCTTTAGAGGAAAACAAATTTCTTTGAGAAAACGAGTTTTCAGAG 990
Db      350 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 369
QY      991 TATCAGCGTTTTCAGTTATGCGAGAAACACACAGATAACGCTTTCACCTTGGATGCGAGAT 1050
Db      370 TyrGlnArgMetGlyValMetSerAsnSerThrGluAsnSerPheThrLeuAspAlaAsp 389
QY      1051 TTT 1053
Db      390 Phe 390

```

RESULT 4

```

S27153
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - golden hamster
N:Alternate names: ribonucleotide reductase small chain
C:Species: Mesocricetus auratus (golden hamster)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S27153
R:Chaudhuri, M.M.; Tonin, P.N.; Srinivasan, P.R.
Biochim. Biophys. Acta 1171, 117-121, 1992
A:Title: cDNA sequence of the small subunit of the hamster ribonucleotide reductase.
A:Reference number: S27153; MUID:93042000; PMID:1384717

```

A:Accession: S27153
A:Molecule type: mRNA
A:Residues: 1-386 <CHA>
C:Cross-references: UNIPROT:Q60561; EMBL:X68127; NID:g49649; PIDN:CAA48232.1; PID:g49650
C:Superfamily: ribonucleoside reductase small subunit
C:Keywords: oxido-reductase

Alignment Scores:
Pred. No.: 6.56e-121 Length: 386
Score: 1421.00 Matches: 266
Percent Similarity: 91.57% Conservative: 38
Best Local Similarity: 80.12% Mismatches: 20
Query Match: 76.98% Indels: 8
DB: 2 Gaps: 2

US-10-698-228-2 (1-1053) x S27153 (1-386)

```
QY 70 GAAAGTGAATAAAGTCAAAAT-----GAAAGCCACTCTTAAAGAAAGATTCT 117
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 59 GluSerLysValSerThrAsnProSerValGluAspGluProLeuLeuArgGluAsnPro 78
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 118 CGCGGTTTCTCATCTTCCATCCAGTACCCTGATATTGGAAATGATATAACAGGCA 177
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 79 ArgArgPheValValPheProIleGluIleHisAspIleTrpLysMetTrpLysLysAla 98
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 178 CAGGCTTCTTCTGGACAGCAGAGAGTTCGACTTATCAAGGATCTCCCTCACTCGAAC 237
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 99 GluAlaSerPheTrpThrAlaGluValAspLeuSerLysAspIleGlnHisTrpGlu 118
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 238 AAGCTTAAAGCAGATGAGAGTACTTCATCTCTCACATCTTAGCCCTTTTTCGAGCCAGT 297
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 119 AlaLeuLysProAspGluArgHisPheIleSerHisValLeuAlaPhePheAlaAlaSer 138
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 298 GATGGAATTTGTAATCAAAATTTGGTGGAGCCCTTTAGTCAGAGGTGAGGTTCCAGAG 357
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 139 AspGlyIleValAsnGluAsnLeuValGluValArgPheSerGlnGluValGlnThrGlu 158
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 358 GCTCGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATGTTCTCACTCAGAGATGACAGT 417
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 159 AlaArgCysPheTrpGlyPheGlnIleAlaMetGluAsnIleHisSerGluMetTrpSer 178
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 418 TTGCTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTATGCAATT 477
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 179 LeuLeuIleAspThrTrpIleLysAspSerLysGluArgGluTrpLeuPheAsnAlaIle 198
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 478 GAAACATCCCTATGTTTAAAGAAAGAGATTTGGGCCCTTGGATGGATGAGATAGAGA 537
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 199 GluThrMetProCysValLysLysLysAlaAspTrpAlaLeuArgTrpIleGlyAspLys 218
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 538 AAATCTACTTTTGGGAAAGAGTGGTGGCTTGTGCTGTAGAGAGTCTTCTCTCA 597
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 219 GluAlaThrTrpGlyGluArgValAlaPheAlaValGluGluIlePhePheSer 238
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 598 GGATCTTTTTCCTGCTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCAGGACTCACTTTT 657
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 239 GlySerPheAlaSerIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPhe 258
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 658 TCAATGAATCATCAGCAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTTC 717
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 259 SerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPhe 278
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 718 CAATACTTGTAAATAGCCTTCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTC 777
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 279 LysHisLeuValHisLysProSerGluGlnArgValGlnGluIleLeuThrAsnAlaVal 298
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 778 AAAATGACAGAGGATTTTAAAGAGGCTTGCAGATTTGGCCTCATTTGGAATGATTC 837
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 299 ArgIleGluGlnGluPheLeuThrGluAlaLeuProValLysLeuIleGlyMetAsnCys 318
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 838 ATTTTGTGAAACAGTACATGAGTTTGTAGCTGACAGATTAATCTGTGGAATTTGGATTC 897
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 319 ThrLeuMetLysGlnTrpIleGluPheValAlaAspArgLeuMetLeuGluLeuGlyPhe 338
```

```
QY 898 TCAAGGTTTTCAGGCAGAAAAATCCTTTTATGATTTTATGAAAAACATTTCTTTAAGAGA 957
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 339 AsnLysIlePheLysValGluAsnProPheAspPheMetGluAsnIleSerLeuGluGly 358
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 958 AAAACAAAATTTCTTTGAGAAACAGATTTCAGAGTATCAGCGTTTTCAGATTATGCAGAA 1017
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 359 LysThrAsnPhePheGluLysArgValGlyGluTrpGlnArgMetGlyValMetSer--- 377
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 1018 ACCACAGATAACGCTCTTCACTTGGATGCAGATTTT 1053
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 378 -----AsnSerPheThrLeuAspAlaAspPhe 386
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
RESULT 5
S24585
ribonucleoside-diphosphate reductase (EC 1.17.4.1) chain M2 - Atlantic surf clam
N;Alternate names: ibonucleoside-diphosphate reductase small chain
C;Species: Spisula solidissima (Atlantic surf clam)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S24585
R;Hunt, T.; Standart, N.
submitted to the EMBL Data Library, November 1990
A;Reference number: S24584
A;Accession: S24585
A;Molecule type: mRNA
A;Residues: 1-384 <HUN>
A;Cross-references: UNIPROT:P07201; EMBL:X55125; NID:g10341; PIDN:CAA38919.1; PID:g10342
C;Superfamily: ribonucleoside reductase small subunit
C;Keywords: oxido-reductase
```

Alignment Scores:
Pred. No.: 9.93e-113 Length: 384
Score: 1331.00 Matches: 257
Percent Similarity: 85.26% Conservative: 38
Best Local Similarity: 74.28% Mismatches: 41
Query Match: 72.10% Indels: 10
DB: 2 Gaps: 4

US-10-698-228-2 (1-1053) x S24585 (1-384)

```
QY 46 GAGAGATCATCTTCAGACACCAAC---GAAAGTGAATAAAG----- 84
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 39 GlnArgSerThrGlnLysTrpProLeuLysGlnGluIleLysProValLysLysSer 58
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 85 TCAATGAAGACAGACCTCTTAAGAAAGAGTCTCGCGGTTTGTCACTTTTCCAAATCCAG 144
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 59 GlnGlnValGluProLeuLeuAlaAspAsnProArgPheValLeuProIleGln 78
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 145 TACCTCTGATATTTGGAAAAATGTATAACAGCGACAGGCTTCTCTCGACAGCAGAAGAG 204
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 79 TyrHisAspIleTrpLysMetTrpLysLysAlaGluAlaSerPheTrpThrAlaGluGlu 98
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 205 GTCGACTATCAAGGATCTCCCTCACTGGAAACAGCTTAAAGCAGATGAGAACTACTTC 264
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 99 ValAspLeuSerLysAspMetAlaHisTrpGluSerLeuLysLysGluLysHisPhe 118
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 265 ATCTCTCACATCTTAGCGCTTTTTCAGCCAGTCAGTAATGTAATGAAAATTTGGTG 324
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 119 IleSerHisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuVal 138
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 325 GACGCTTTAGTCAGGAGGTGTCAGGTTTCCAGAGCTCGCTGTTTCTATGGCTTTCAAAT 384
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 139 GluArgPheSerLysGluValGlnValThrGluAlaArgCysPheTrpGlyPheGlnIle 158
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 385 CTCACTCAGAAATTTCACTCAGAGATGTACAGTTTGTGTAGACACTTACATCAGAGAT 444
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 159 AlaMetGluAsnIleHisSerGluMetTrpSerLeuLeuIleAspThrTrpIleLysAsp 178
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 445 CCCAAGAAAAGGAATTTTATTATGCAATTCGAATTCGAACCATGCCCTATGTTAAGAAAAA 504
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 179 ProGlnGluArgAspPheLeuPheAsnAlaIleGluThrMetProCysValLysGluLys 198
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 505 GCAGATTGGGCTTTCGATGATGATAGACATAGAAAAATCTACTTTTGGGAAAGAGTGGTG 564
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
```

Db 199 AlaAepTrpAlaMetArgTrpIleAsnAepAepSerSerSerTyrAlaGluArgValVal 218
 QY 565 GCCTTGGCTGCTAGAACAGATTTCTTCTCAGGATCTTTGCTGCTATATTTCTGGCTA 624
 Db 219 AlaPheAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeu 238
 QY 625 AAGAAGAGAGGCTTTATGCGCAGGACTCCTCTTTTCCAATGAATCATCAGCAGAGATGAA 684
 Db 239 LysLysArgGlyMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAepGlu 258
 QY 685 GCACTTCAGTGTACTTCTGCTGCTGATGTTCCCAATACCTAGTAATTAAGCCTTCAGAA 744
 Db 259 GlyLeuHisCysAepPheAlaCysLeuMetPheSerHisLeuValAsnLysProSerGln 278
 QY 745 GAAAGGTCAGGAGATCATTTGATGCTGCTCAAAATTCAGCAGGAGTTTAAACAGAA 804
 Db 279 GluArgIleHisGlnIleLeuAepGluAlaValLysIleGluInValPheLeuThrGlu 298
 QY 805 GCCTTGCAGTTGGCTCATTTGGAATGAATTTGATTTGATGAACAGTACATTTGAGTTT 864
 Db 299 AlaLeuProCysArgLeuIleGlyMetAsnCysAepLeuMetArgGlnTyrIleGluPhe 318
 QY 865 GTAGCTGACAGATTACTTGTGGAATCTGATTTCAAGGTTTTCAGCAGAGAAATCCT 924
 Db 319 ValAlaAepArgLeuLeuGluLeuLysCysAepLysLeuTyrAsnLysGluAsnPro 338
 QY 925 TTTGATTTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTT 984
 Db 339 PheAepPheMetGluHisIleSerLeuGluGlyLysThrAsnPhePheGluLysArgVal 358
 QY 985 TCAGAGTATCAGGCTTTTTCAGTTATGGCA-----GAAACACACAGAT--AACGCTTC 1035
 Db 359 GlyGluTyrGlnLysMetGlyValMetSerGlyGlyAsnThrGlyAepSerHisAlaPhe 378
 QY 1036 ACCTTGGATGCAGATTTT 1053
 Db 379 ThrLeuAepAlaAepPhe 384

RESULT 6
 T30782
 ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - vaccinia virus (strain
 N) Alternate names: ribonucleotide reductase, small subunit
 C:Species: vaccinia virus
 A:Variety: strain Ankara
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 R:Antoine, G.; Scheifflinger, F.; Falkner, F.G.; Dorner, F.
 submitted to the EMBL Data Library, March 1997
 A:Description: The complete genomic sequence of the Modified Vaccinia Ankara (MVA) strain
 A:Reference number: Z20877
 A:Accession: T30782
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-319 <ANT>
 A:Cross-references: UNIPROT:O57175; EMBL:U94848; PIDN:AAB96415.1
 A:Experimental source: strain Ankara
 C:Genetics:
 A:Note: MVA032L
 C:Function:
 A:Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleoside
 C:Superfamily: ribonucleoside reductase small subunit
 C:Keywords: deoxyribonucleotide biosynthesis; oxidoreductase

Alignment Scores:
 Pred. No.: 4,13e-108 Length: 319
 Score: 1280.00 Matches: 239
 Percent Similarity: 85.31% Conservative: 34
 Best Local Similarity: 74.69% Mismatches: 45
 Query Match: 69.34% Indels: 2
 DB: 2 Gaps: 2

US-10-698-228-2 (1-1053) x T30782 (1-319)

QY 94 GAGCCACTCCTAAGAAAGAGTTCTCGCCGGTTTGTCTCATCTTTCCAATCCAGTACCCCTGAT 153
 Db 2 GluProIleLeuAlaProAsnProAsnArgPheValIlePheProIleGlnTyrHisAep 21
 QY 154 ATTTGAAATATGATATAACAGGACAGCTTCTTCTGACAGCAGCAGAGAGGTGACTTA 213
 Db 22 IleTrpAsnMetTyrLysLysAlaGluAlaSerPheTrpValGluGluValAaspIle 41
 QY 214 TCAAGAGATCTCCCTCCTCAGTGAACAAGCTTAAAGCAGATGAGAGTAGTCTTCATCTCTCAC 273
 Db 42 SerLysAaspIleAsnAepTrpAsnLysLeuThrProAepGluLysTyrPheIleLysHis 61
 QY 274 ATCTTAGCCCTTTTTCAGCCAGTGTGAATGTAATGAAATTTGGTGGAGCGCTTT 333
 Db 62 ValLeuAlaPhePheAlaSerAepGlyIleValAsnGluAsnLeuAlaGluArgPhe 81
 QY 334 AGTCAGAGGTGCGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTTCTCATCAG 393
 Db 82 CysThrGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnMetAlaIleGlu 101
 QY 394 AATGTTCACTCAGAGATGTACAGTTGCTGATAGACATTTACATCAGAGATCCCAAGAA 453
 Db 102 AsnIleHisSerGluMetTyrSerLeuLeuIleAaspThrTyrValLysAepSerAsnGlu 121
 QY 454 AGCGAATTTTATTAATGCAATTGAAACCATGCCCTATGTTAAGAAAAAGACAGATGG 513
 Db 122 LysAsnTyrLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAaspTrp 141
 QY 514 GCCTTCCGATGGATAGACAGATGAAATCTACTTTTCGGGAAAGAGTGGTGGCTTGTCT 573
 Db 142 AlaGlnLysTrpIleHisAep--SerAlaGlyTyrGlyGluArgLeuIleAlaPheAla 160
 QY 574 GCTGTAGAGAGGATTTTCTTCTCAGGATCTTTGCTGCTATATTTCTGGCTAAAGAGAGA 633
 Db 161 AlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysArg 180
 QY 634 GGTCTTATGCCAGGACTCACTTTTCCAATGAATCATCAGCAGAGATGAGGACTTCAC 693
 Db 181 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAepGluGlyLeuHis 200
 QY 694 TGTGACTTTGCTTGCCTGATGTTTCCAAATCTTAAATTAAGCCTTCAGAGAAAGGGTCT 753
 Db 201 CysAepPheAlaCysLeuMetPheLysHisLeuHisProProSerGluGluThrVal 220
 QY 754 AGGAGATCATTTGATGCTGTCAAAATTTAGCAGAGGATTTTAAACAGAAAGCTTGCCA 813
 Db 221 ArgSerIleIleThrAaspAlaValSerIleGluGlnGluPheLeuThrAlaAlaLeuPro 240
 QY 814 GTTGGCTCATTTGGAATGATTCATTTGATGAACAGTACATTTGATTTGTAGCTGAC 873
 Db 241 ValLysLeuIleGlyMetAsnCysGluMetLysThrTyrIleGluPheValAlaAasp 260
 QY 874 AGATTACTTGTGGAACCTTGATTTCTCAAGATTTTTCAGGCAGAGAAATCTTTTGATTTT 933
 Db 261 ArgLeuIleSerGluLeuGlyPheLysLysIleTyrAsnValThrAsnProPheAaspPhe 280
 QY 934 ATGGAACCATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACAGATTTTCAGAGPAT 993
 Db 281 MetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGluTyr 300
 QY 994 CAGCGTTTTCAGTTATGGCAGAAACACAGATTAACGCTTTCACCTTGGATGACAGATTTT 1053
 Db 301 GlnLysMetGlyValMetSerGln--GluAaspAsnHisPheSerLeuAepValAaspPhe 319

RESULT 7
 B72154
 E4L protein - variola minor virus (strain Garcia-1966)

C:Species: variola minor virus
 C:Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 09-Jul-2004
 C:Accession: B72154
 R:Shchelkunov, S.N.; Totmenin, A.V.; Gutorov, V.V.; Safronov, P.F.; Massung, R.F.; Lopar
 submitted to GenBank, March 1998
 A:Description: Analysis of the complete coding sequence of DNA of alastrim variola minor

A:Reference number: A72150
A:Accession: B72154
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-319 <SHC>
A:Cross-references: UNIPROT:Q89087; GB:Y16780; NID:G5830555; PTDN:CA54628.1; PTD:G58305
A:Experimental source: strain Garcia-1966
C:Genetics:
A:Gene: E4L
C:Superfamily: ribonucleoside reductase small subunit

Alignment Scores:
Pred. No.: 4.13e-108 Length: 319
Score: 1280.00 Matches: 240
Percent Similarity: 85.31% Conservative: 33
Best Local Similarity: 75.00% Mismatches: 45
Query Match: 69.34% Indels: 2
DB: 2 Gaps: 2

US-10-698-228-2 (1-1053) x B72154 (1-319)

QY 94 GAGCCACTCTTAAGAAAGATTCTCGCGTTTGTCTATCTTCCATCCAGTACCTGTAT 153
Db 2 GluProIleLeuAlaLysAsnProAsnArgPheValIlePheProIleGlnTyrHisAsp 21

QY 154 ATTTCGAAATCTATAACAGCAGCAGGCTTCTCTGACAGCAGAGAGGTCCACTTA 213
Db 22 IleTrpAsnMetTyrLysLysAlaGluAlaSerPheTrpThrValGluGluValAspIle 41

QY 214 TCAAGGATCTCCCTCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 273
Db 42 SerLysAspIleAsnAspTrpAsnLysLeuThrProAspGluLysTyrPheIleLysHis 61

QY 274 ATCTTAGCTTTTTCGACGAGTGAATGGAATGGAATGGAATGGAATGGAATGGAATG 333
Db 62 ValLeuAlaPheAlaAlaSerAspGlyIleValAsnGluAsnLeuAlaGluArgPhe 81

QY 334 AGTCAGGAGGTTCAGAGGCTCGCTGTTCTATGGCTTTCATATTCATATTCATATTC 393
Db 82 CysIleGluValGluIleThrGluAlaArgCysPheTyrGlyPheGlnMetAlaIleGlu 101

QY 394 ATGTTCAGTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 453
Db 102 AsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrValLysAspSerAsnGlu 121

QY 454 AGGGAATTTTATTAATCAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTC 513
Db 122 LysAsnTyrLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAspTrp 141

QY 514 GCCTTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 573
Db 142 AlaGlnLysTrpIleHisasp--SerAlaGlyTyrGlyGluArgLeuIleAlaPheAla 160

QY 574 GCTGTAGAGAGATTTTCTCTCAGGATCTTTCGCTGATATTCCTGCTAAGAGAGA 633
Db 161 AlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArg 180

QY 634 GGTCTATGCCAGGATCTACTTTTCCAAATGAACTCATCAGCAGAGATGAAGACTTCAC 693
Db 181 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHis 200

QY 694 TGTGACTTGTCTGCTGATGTTTCCAAATCTAGTAATTAAGCCCTTCAGAAAGGGTTC 753
Db 201 CysAspPheAlaCysLeuMetPheLysHisLeuLeuTyrProProSerGluGluThrVal 220

QY 754 AGGAGATCATTTGTTGATCTGCTGCTCAAAATTTAGCAGGAGTCTTTTAAACAGAGCTTGCCA 813
Db 221 ArgSerIleIleThrAspAlaValSerIleGluGlnGluPheLeuThrValAlaLeuPro 240

QY 814 GTTGGCTCATTTGGAATGATTTGCATTTTGTGATGAACAGTACATGAGTTTGTAGCTGAC 873
Db 241 ValLysLeuIleGlyMetAsnCysGluMetMetLysThrTyrIleGluPheValAlaAsp 260

QY 874 AGATTACTTGTGAACCTTGATTTCTCAAGGTTTTCAGCGAGAAATCCTTTTGATTTT 933
Db 261 ArgLeuIleSerGluLeuGlyPheLysLysIleTyrAsnValThrAsnProPheAspPhe 280

QY 934 ATGGAACATTTCTTTAGAGAAACAAACAAATTTCTTTGAGAACGAGTTTCAGACTAT 993
Db 281 MetGluAsnIleSerLeuGluGlyLysThrAsnPheGluLysArgValGlyGluTyr 300

QY 994 CAGCGTTTTCAGTTATGCGAGAAACACAGATAAAGCTCTTCACCTTGATGACAGATTTT 1053
Db 301 GlnLysMetGlyValMetSerGln---GluAspAsnHisPheSerLeuAspValAspPhe 319

RESULT 8
RDVZVV
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - vaccinia virus
N/Alternate names: P4L protein
C/Species: vaccinia virus
C/Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
C/Accession: A29892; I42506; F36213
R/Slabaugh, M.; Roseman, N.; Davis, R.; Mathews, C.
J. Virol. 62, 519-527, 1988
A/Title: Vaccinia virus-encoded ribonucleotide reductase: sequence conservation of the
A/Reference number: A29892; MUID:88091062; PMID:2826813
A/Accession: A29892
A:Molecule type: DNA
A:Residues: 1-319 <SLA>
A:Cross-references: UNIPROT:P11158; GB:M19117; NID:G335808; PTDN:AAA88680.1; PID:G335809
A:Experimental source: strain WR
R/Goebel, S.J.; Johnson, G.P.; Perkus, M.E.; Davis, S.W.; Winslow, J.P.; Paoletti, E.
Virology 179, 517-563, 1990
A/Title: Appendix to "The complete DNA sequence of vaccinia virus".
A/Reference number: A42501
A:Accession: I42506
A:Molecule type: DNA
A:Residues: 1-212, 'Y', 214-319 <GOR>
A:Cross-references: GB:M35027; NID:G335317; PTDN:AAA48018.1; PID:G335366
A:Experimental source: strain Copenhagen
R/Goebel, S.J.; Johnson, G.P.; Perkus, M.E.; Davis, S.W.; Winslow, J.P.; Paoletti, E.
Virology 179, 247-266, 1990
A/Title: The complete DNA sequence of vaccinia virus.
A/Reference number: A42531; MUID:91021027; PMID:2219722
A/Contents: annotation; possible protein-coding frames
A/Note: neither amino acid nor nucleotide sequence is given
R/Roseman, N.A.; Slabaugh, M.B.
Virology 178, 410-418, 1990
A/Title: The vaccinia virus HindIII F fragment: nucleotide sequence of the left 6.2 kb.
A/Reference number: A36213; MUID:91020979; PMID:2219701
A:Accession: F36213
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-319 <ROS>
A:Cross-references: EMBL:M34368; NID:G335618; PTDN:AAA48244.1; PID:G335624
A:Experimental source: strain WR
C/Function:
A/Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleos
C/Superfamily: ribonucleoside reductase small subunit
C/Keywords: DNA replication; iron; metalloprotein; oxidoreductase
F/70,101,104,163,197,200/Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, His) #sta
F/108/Active site: Tyr (stable tyrosyl radical) #status predicted

Alignment Scores:
Pred. No.: 5.09e-108 Length: 319
Score: 1279.00 Matches: 239
Percent Similarity: 85.31% Conservative: 34
Best Local Similarity: 74.69% Mismatches: 45
Query Match: 69.28% Indels: 2
DB: 1 Gaps: 2

US-10-698-228-2 (1-1053) x RDVZVV (1-319)

QY 94 GAGCCACTCTTAAGAAAGATTCTCGCGTTTGTCTATCTTCCATCCAGTACCTGTAT 153
Db 2 GluProIleLeuAlaProAsnProAsnArgPheValIlePheProIleGlnTyrTyrAsp 21


```
QY 154 ATTGGAATATGTATAACAGGACAGGCTTCTTCTGGACAGCAGAGAGGTTCGACTTA 213
DB 22 IletPAsnMetTyrLysLysAlaGluAAsrPheThrValGluGluValAspIle 41
QY 214 TCAAGGATCTCCCTCAGTGAACAGCTTAAGACAGATGAGAGTACTTCTCTCAC 273
DB 42 SerLysAspIleAsnAspTrpAsnLysLeuThrProAspGluLysTyrPheIleLysHis 61
QY 274 ATCTTAGCCTTTTTCACGACGATGGAATGTAAATTCGAGGAGGCTTT 333
DB 62 ValLeuAlaPhePheAlaSerAspGlyIleValAsnGluAsnLeuAlaGluArgPhe 81
QY 334 AGTCAGGAGGTTCAGAGGCTCGCTGTTCTATGGCTTTCATAATTCATCTCAG 393
DB 82 CysThrGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnMetAlaIleGlu 101
QY 394 ATGTTCTACTAGAGATGACATGATGCTGATAGACATCTATACATCAGAGATCCCAAGAA 453
DB 102 AsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrValLysAspSerAsnGlu 121
QY 454 AGGGAATTTTATTAATGCAATGAAACCATGCTCTATGTTAAGAAAAGCAGATTGG 513
DB 122 LysAsnTyrLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAspTrp 141
QY 514 GCCTTGCAGATGACATAGAAATCTACTTTTGGGGAAGAGCTGGCTTTGCT 573
DB 142 AlaGlnLysTrpIleHisasp--SerAlaGlyTyrGlyLysGluLeuAlaPheAla 160
QY 574 GCTGTAGAGGAGTCTTCTCTCAGGATCTTTGCTGCTATATCTGCTAAGAGAGA 633
DB 161 AlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArg 180
QY 634 GGTCTTATGCCAGGACTCACTTTTCCAAATGAACATCATCAGCAGAGATGAGGACTTCAC 693
DB 181 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHis 200
QY 694 TGTGACTTGTGCTGCTGATGTTCCAACTACTAGTAAATAGCCCTCAGAAGAGGCTC 753
DB 201 CysAspPheAlaCysLeuMetPheLysHisLeuLeuIleProProSerGluGluThrVal 220
QY 754 AGGAGATCATTTGTCATCTCTCAAAATGAGCAGGAGTCTTAAACAGAGCTTGC 813
DB 221 ArgSerIleIleThrAspAlaValSerIleGluGlnPheLeuThrAlaAlaLeuPro 240
QY 814 GTTGCCCTCATGGAATGAAATTCATTTGATGAACACATACATTGAGTTGTGACTGAC 873
DB 241 ValLysLeuIleGlyMetAsnGlyMetMetLysThrTyrIleGluPheValAlaAsp 260
QY 874 AGATTACTTGTGGAATCTGATCTCAAGGTTTTCAGCGAATAATCTTTGATTTT 933
DB 261 ArgLeuIleSerGluLeuGlyPheLysLysIleTyrAsnValThrAsnProPheAspPhe 280
QY 934 ATGGAACATTTCTTGAAGGAAACAAATTTCTTGAAGAACGAGTCTCAGAGTAT 993
DB 281 MetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGluTyr 300
QY 994 CAGCGCTTTGCAAGTTATGCGAAGAACCCAGATAACGCTTCCACCTTGGATCGAGATTTT 1053
DB 301 GlnLysMetGlyValMetSerGln--GluAspAsnHisPheSerLeuAspValAspPhe 319

RESULT 9
T28466
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - variola major virus
C:Species: variola major virus
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 15-Sep-2003
C:Accession: T28466
R:Massung, R.F.; Eposito, J.J.; Liu, L.I.; Qi, J.; Utterback, T.R.; Knight, J.C.; Aubin
Nature 366, 748-751, 1993
A:Title: Potential virulence determinants in terminal regions of variola smallpox virus
A:Reference number: 220488; MUID:94088747; PMID:8264798
A:Accession: T28466
A>Status: preliminary; translated from GB/EMBL/DBJ
```

```
A:Molecule type: DNA
A:Residues: 1-333 <MAS>
A:Cross-references: EMBL:L22579; NID:9623595; PID:AAA60776.1; PID:9438946
A:Experimental source: strain "Bangladesh-1975"
C:Function:
A:Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleos
C:Superfamily: ribonucleoside reductase small subunit
C:Keywords: deoxyribonucleotide biosynthesis; oxidoreductase

Alignment Scores:
Pred. No.: 5,14e-108 Length: 333
Score: 1279.00 Matches: 239
Percent Similarity: 85.62% Conservative: 35
Best Local Similarity: 74.69% Mismatches: 44
Query Match: 69.28% Indels: 2
DB: 2 Gaps: 2

US-10-698-228-2 (1-1053) x T28466 (1-333)
QY 94 GAGCCACTCTTAAGAAAGAGTCTCGCGGTTTCTCATCTTCCAATCCAGTACCCTGAT 153
DB 16 GluProIleLeuAlaLysAsnProAsnArgPheValIlePheProIleGlnTyrHisasp 35
QY 154 ATTGGAATATGTATAACAGGACAGGCTTCTTCTGGACAGCAGAGAGGTTCGACTTA 213
DB 36 IletPAsnMetTyrLysLysAlaGluAAsrPheThrValGluGluValAspIle 55
QY 214 TCAAGGATCTCCCTCAGTGAACAGCTTAAGACAGATGAGAGTACTTCTCTCAC 273
DB 56 SerLysAspIleAsnAspTrpAsnLysLeuThrProAspGluLysTyrPheIleLysHis 75
QY 274 ATCTTAGCCTTTTTCAGCAGGCTGATGGAATGTAAATTCGAGGAGGCTTT 333
DB 76 ValLeuAlaPhePheAlaSerAspGlyIleValAsnGluAsnLeuAlaGluArgPhe 95
QY 334 AGTCAGGAGGTGAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCATAATTCATCCAG 393
DB 96 CysIleGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnMetAlaIleGlu 115
QY 394 ATGTTCTACTCAGAGATGACATGCTGCTGATGACACTTATCAGAGATCCCAAGAAA 453
DB 116 AsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrValLysAspSerAsnGlu 135
QY 454 AGGGAATTTTATTAATGCAATGAAACCATGCTCTATGTTAAGAAAAGCAGATTGG 513
DB 136 LysAsnTyrLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAspTrp 155
QY 514 GCCTTGCAGATGACATAGAAATCTACTTTTGGGGAAGAGTGTGCGCTTTGCT 573
DB 156 AlaGlnLysTrpIleHisasp--SerAlaSerTyrGlyLysGluArgLeuIleAlaPheAla 174
QY 574 GCTGTAGAGGAGTCTTCTCTCAGGATCTTTGCTGCTATATTCGCTGCTGCTGCTGCT 633
DB 175 AlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArg 194
QY 634 GGTCTTATGCCAGGACTCACTTTTCCAAATGAACATCATCAGCAGAGATGAGGACTTCAC 693
DB 195 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHis 214
QY 694 TGTGACTTGTGCTGCTGATGTTCCAAATCTAGTAAATTAAGCCTTCAGAAGAAAGGCTC 753
DB 215 CysAspPheAlaCysLeuMetPheLysHisLeuLeuIleProProSerGluGluThrVal 234
QY 754 AGGAGATCATTTGCTGCTGCTCAAAATTCAGCAGGAGTCTTTTAAACAGAGCCTTGCCA 813
DB 235 ArgSerIleIleThrAspAlaValSerIleGluGlnPheLeuThrValAlaLeuPro 254
QY 814 GTTGCCCTCATGGAATGAAATTCGATTTTGGATGAACAGTACATGATGATTTGTAGTGCAC 873
DB 255 ValLysLeuIleGlyMetAsnGlyMetMetLysThrTyrIleLysPheValAlaAsp 274
QY 874 AGATTACTTGTGGAATCTGATCTCAAGGTTTTCAGGAGAGATTTTTCAGGAGAGATTTT 933
DB 874 AGATTACTTGTGGAATCTGATCTCAAGGTTTTCAGGAGAGATTTTTCAGGAGAGATTTT 933
```

```
Db 275 ArgLeuIleSerGluLeuGlyPheLysLysIleTyrAsnValThrAsnProPheAspPhe 294
QY 934 ATGGAAAACATTTCTTACAGAGAAAACAAATTTCTTGGAGAAACGAGTTTCAGAGTAT 993
Db 295 MetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGluTyr 314
QY 994 CAGCGTTTTCAGAGTTATGGCAGAAACCCAGATAACGCTCTTCACCTTGGATCGACATTTT 1053
Db 315 GlnLysMetGlyValMetSerGln---GluAspAsnHisPheSerLeuAspValAspPhe 333

RESULT 10
H36839
CBL protein - variola virus (strain India-1967)
C:Species: variola virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: H36839
R:Blinov, V.M.
submitted to GenBank, November 1992
A:Reference number: A36859
A:Accession: H36839
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-333 <BL>
A:Cross-references: UNIPROT:P33799; GB:X69198; NID:G456758; PIDN:CAA48969.1; PID:G297209
C:Superfamily: ribonucleoside reductase small subunit

Alignment Scores:
Pred. No.: 7,81e-107 Length: 333
Score: 1266.00 Matches: 237
Percent Similarity: 85.00% Conservative: 35
Best Local Similarity: 74.06% Mismatches: 46
Query Match: 68.58% Indels: 2
DB: 2 Gaps: 2

US-10-698-228-2 (1-1053) x H36839 (1-333)

QY 94 GAGCCACTCTTAAGAAAGAGTTCTCCGCGTTTGTCTATCTTTCCAATCCAGTACCTCGAT 153
Db 16 GluProIleLeuAlaLysAsnProAsnArgPheValIlePheProIleGlnTyrHisAsp 35
QY 154 APTTGGAAAATATATAACAGGCACAGGCTTCTTCTGGACAGCAGAAAGAGTCGACTTA 213
Db 36 IleTyrAsnMetTyrLysLysAlaGluAlaSerPheTyrThrValGluGluValAspIle 55
QY 214 TCAAGAGTCTCCCTCAGTGGACAGCTTAAGACAGATGAGAGTACTTCTCTCTAC 273
Db 56 SerLysAspIleAsnAspTrpAsnLysLeuThrProAspGluLysTyrPheIleLysHis 75
QY 274 ATCTTAGCCTTTTTCAGCCAGTGTAAATGTAATGAAATTTGGTGAGCGCTTT 333
Db 76 ValLeuAlaPheAlaAlaSerAspGlyIleValAsnGluAsnLeuAlaGluArgPhe 95
QY 334 ACTCAGGAGGTCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAAATTCATCGAG 393
Db 96 CysIleGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnMetAlaIleGlu 115
QY 394 ATGTTCTACTCAGAGATGTACAGTTGCTGATAGACACTTACATCAGATCCCAAGAA 453
Db 116 AsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrValLysAspSerAsnGlu 135
QY 454 AGGGAATTTTATTAATCAATGAAACCATGCGCTATGTTTAAAGAAAACACAGATTGS 513
Db 136 LysAsnTyrLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAspTrp 155
QY 514 GCCTTGCGATGATGACAGATAAATCTACTTTTGGGAAAGAGTGTGGCTTTTGCT 573
Db 156 AlaGlnLysTrpIleHisAsp---SerAlaGlyTyrGlyGluArgLeuIleAlaPheAla 174
QY 574 GCTGTAGAGAGGATTTTCTTCAGAGATCTTTGCTGCTATATTCGTGGCTTAAGAGAGA 633
Db 175 AlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArg 194
QY 634 GGTCTTATCCAGGACTCACTTTTCCATGAACTCATCAGCAGAGATGAAGGACTTCAC 693
```

```
195 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHis 214
QY 694 TGTGACTTTCCTTCCTGCTGATGTTCCAACTACTAGTAAATAGCCTTCAGAGAAAGGTC 753
Db 215 CysAspPheAlaCysLeuMetPheLysHisLeuTyrProProSerGluGluThrVal 234
QY 754 AGGGAGATCATTTGCTGATGCTGTCAAAATTCAGCAGGAGTTTAAACAGAACCTTGGCA 813
Db 235 ArgSerIleIleThrAspAlaValSerIleGluGlnGluPheLeuThrValAlaLeuPro 254
QY 814 GTTGGCCTCATTTGGAATGAATTCATTTTGTGATGAACAGTACATAGTTGTAGCTGAC 873
Db 255 ValLysLeuIleGlyMetAsnCysGluMetMetLysThrTyrMetGluPheValAlaAsp 274
QY 874 AGATTACTTGTGGAACCTTGGATTTCTCAAAGGTTTTCAGGCAGAAAATCCTTTGATT 933
Db 275 ArgLeuIleSerGluLeuGlyPheLysArgIleTyrAsnValThrAsnProSerAspPhe 294
QY 934 ATGGAACCAATTTCTTTAGAGGAAAACAAATTTCTTTCAGAAAACGAGTTTCAGAGTAT 993
Db 295 MetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGluTyr 314
QY 994 CAGCGTTTTCAGTATTAGCAGAACCCAGATAACGCTTTCACCTTGGATGCGAGATTT 1053
Db 315 GlnLysMetGlyValMetSerGln---GluAspAsnHisPheSerLeuAspValAspPhe 333

RESULT 11
T30470
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - Lymantria dispar nuclear
N:Alternate names: ribonucleotide reductase small subunit homolog
C:Species: Lymantria dispar nuclear polyhedrosis virus, LdMNPV
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T30470
R:Kuzio, J.; Pearson, M.N.; Harwood, S.H.; Funk, C.J.; Evans, J.T.; Slavicek, J.M.; Rohrer
Virolgy 253, 17-34, 1999
A:Title: Sequence and analysis of the genome of a baculovirus pathogenic for Lymantria
A:Reference number: Z20836; MUID:99124785; PMID:9887315
A:Accession: T30470
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-348 <KU>
A:Cross-references: UNIPROT:Q9YMK7; EMBL:AF081810; PIDN:AAC70306.1
C:Function:
A:Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleoside
C:Superfamily: ribonucleoside reductase small subunit
C:Keywords: oxidoreductase; pyrimidine deoxynucleotide metabolism

Alignment Scores:
Pred. No.: 3,78e-106 Length: 348
Score: 1258.50 Matches: 246
Percent Similarity: 80.63% Conservative: 37
Best Local Similarity: 70.09% Mismatches: 59
Query Match: 68.17% Indels: 9
DB: 2 Gaps: 3

US-10-698-228-2 (1-1053) x T30470 (1-348)

QY 1 ATGGCGCACCCGGAAGGCGGAGCGCGCGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 7 MetProAlaProGluArgPro-----PheAspProArgAlaProAla 20
QY 61 GACACCAACGAAGTAGTAATTAAGTCAATTAAGAGAGCACTCTCTAAGAAAGAGTTCTCGC 120
Db 21 Pro-----GluArgProPheAspProArgAlaGluProLeuLeuArgGluAsnProArg 38
QY 121 CGTTTGTTCATCTTCCATCCAGTACCTCCGATATTTGGAAATGTAATAACAGCACAG 180
Db 39 ArgPheValIlePheProIleGlnTyrProAspMetTrpArgMetTyrLysLysAlaGlu 58
QY 181 GCTTCCTCTTCGACAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGAAACAA 240
Db 59 AlaSerPheTrpThrValGluValAspLeuSerLysAspThrSerAspTrpGluArg 78
```



```
QY 1003 GCAGTTATGCGCA-----GAAACCCAGAGT---AAGCTCTTCCACCTTGATGCGAGATTTT 1053
Db 280 GlyValMetSerGlyGlyAsnThrGlyAspSerHisAlaPheThrLeuAspAlaAspPhe 299
RESULT 13
T18876
hypotheical protein C03C10.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18876
R:Derks, M.
submitted to the EMBL Data Library, August 1994
A:Reference number: Z19036
A:Accession: T18876
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-381 <WIL>
A:Cross-references: UNIPROT:P42170; EMBL:Z35637; PIDN:CAA84688.1; GSPDB:GN00021; CESP:CO
A:Experimental source: clone C03C10
C:Genetics:
A:Gene: CESP.C03C10.3
A:Map position: 3
A:Introns: 18/3; 137/3; 256/2
C:Superfamily: ribonucleoside reductase small subunit
Alignment Scores:
Pred. No.: 4 27e-102 Length: 381
Score: 1214.00 Matches: 231
Percent Similarity: 81.10% Conservative: 48
Best Local Similarity: 67.15% Mismatches: 61
Query Match: 65.76% Indels: 4
DB: 2 Gaps: 2
US-10-698-228-2 (1-1053) x T18876 (1-381)
QY 34 CTGGATCAGGTAGGAGATCATCTTCAGAC-----ACCAAGAAAGTGAATAAAG--- 84
Db 38 ValAspGlnThrLysAlaAlaSerAlaGluThrAsnAsnGluSerGluValAsnGlu 57
QY 85 ----TCAATGAAGACCACTCTCTAGAAAGAGTCTCGCGGTTGTCTATCTTTTCAATC 141
Db 58 LeuAspAlaAspGluProMetLeuGlnAspLeuAspAsnArgPheValIlePheProLeu 77
QY 142 CAGTACCTTGATTTGGAAATGTATAAACAGGACAGCTTCCTTCGGACAGAGAA 201
Db 78 LysHisHisAspIleTrpAsnPheTyrLysLysAlaValAlaSerPheTrpThrValGlu 97
QY 202 GAGTGCAGCTTATCAAGAGTCTCCCTCACTCGAAACAAGCTTAAAGCAGATGAGAAGTAC 261
Db 98 GluValAspLeuGlyLysAspMetAsnAspTrpGluLysMetAsnGlyAspGluGlnTyr 117
QY 262 TTCATCTCTCATCTTAGCCCTTTTTCAGCAGATGGAATTTGAAATGAAATTTG 321
Db 118 PheIleSerArgIleLeuAlaPhePheAlaSerAspGlyLeuValAsnGluAsnLeu 137
QY 322 GTGGAGCGCTTTAGTCAGAGGTGCGGTTCAGAGGCTCGCTGTTCTATGCTTTCAA 381
Db 138 CysGluArgPheSerAsnGluValGlnValSerGluAlaArgPheTyrGlyPheGln 157
QY 382 ATCTCTCATCGAATGTTTCACATCAGAGATGTATACAGTTTCTGATACACTTACATCAGA 441
Db 158 IleAlaIleGluAsnIleHisSerGluMetTyrSerLysLeuIleGluThrTyrIleArg 177
QY 442 GATCCCAAGAAAGGAATTTTTTATTAATGCAATTTGAAACCATCGCCCTATGTTAAGAA 501
Db 178 AspGluThrGluArgAsnThrLeuPheAsnAlaValAspGluPheGluPheIleLysLys 197
QY 502 AAAGCAGATTGGCCCTTGGCATGGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTG 561
Db 198 LysAlaAspTrpAlaLeuArgTrpIleSerAspLysLysAlaSerPheAlaGluArgLeu 217
QY 562 GTGGCCTTGTGCTGTAGAACGAGTTTCTCTCAGGATCTTTTCTGCTATATTTCTGG 621
```

```
Db 218 IleAlaPheAlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrp 237
QY 622 CTAAGAAGAGAGAGTCTTATGCCAGGACTCAGCTTTTCCAAATGAACTCATCAGCAGAGAT 681
Db 238 LeuLysArgGlyLeuMetProGlyLeuThrHisSerAsnGluLeuIleSerArgAsp 257
QY 682 GAAGGAGCTTCACTGTGCTACTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 741
Db 258 GluGlyLeuHisArgAspPheAlaCysLeuLeuTyrSerLysLeuGlnLysLysLeuThr 277
QY 742 GAAGAAGGTCAGGAGATCATTTGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 801
Db 278 GlnGlnArgIleTyrAspIleIleLysAspAlaValAlaIleGluGlnGluPheLeuThr 297
QY 802 GAAGCCTTGCAGTTCGCCTCATTTGGAATGAATGCAATTTTGTGAAACACATGATTCGAG 861
Db 298 GluAlaLeuProValAspMetIleGlyMetAsnCysArgLeuMetSerGlnTyrIleGlu 317
QY 862 TTTGTAGCTGACAGATTAATCTTGTGGAATCTGGAATCTCAAAGGTTTTTTCAGCGCAGAAAT 921
Db 318 PheValAlaAspHisLeuLeuValGluLeuGlyCysAspLysLeuTyrLysSerLysAsn 337
QY 922 CTTTGTGATTTATGAAAAACATTTCTTTAGAGAGAAAAACAAATTTCTTTGAGAAACGA 981
Db 338 ProPheAspPheMetGluAsnIleSerIleAspGlyLysThrAsnPheGluLysArg 357
QY 982 GTTTCAGATTAATCAGCGTTTTCAGATTAATGCGAAGAACACAGATTAAGCTTTCACCTTG 1041
Db 358 ValSerGluTyrGlnArgProGlyValMetValAspGluAlaGluArgGlnPheAspLeu 377
QY 1042 GATCAGAGATTTT 1053
Db 378 GluAlaAspPhe 381
RESULT 14
A26916
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein J1271; protein YJL026w; ribonucleotide reductase small chain
C:Species: Saccharomyces cerevisiae
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: A26916; S07605; S56798
R:Elledge, S.J.; Davis, R.W.
Mol. Cell. Biol. 7, 2783-2793, 1987
A:Title: Identification and isolation of the gene encoding the small subunit of ribonucleoside-diphosphate reductase
A:Reference number: A26916; MUID:88038817; PMID:3313004
A:Molecule type: DNA
A:Residues: 1-399 <ELL>
A:Cross-references: UNIPROT:P09938; EMBL:M17221
A>Note: the authors translated the codon GAA for residue 101 as Lys, GCT for residue 102 as Ser, GCT for residue 107 as Phe, and GAA for residue 108 as Trp
A>Note: the sequence shown follows the authors' translation
R:Hurd, H.K.; Roberts, C.W.; Roberts, J.W.
Mol. Cell. Biol. 7, 3673-3677, 1987
A:Title: Identification of the gene for the yeast ribonucleotide reductase small subunit
A:Reference number: S07605; MUID:88065506; PMID:3316984
A:Accession: S07605
A:Molecule type: DNA
A:Residues: 1-399 <HUR>
A:Cross-references: EMBL:M17789; MID:g172449; PIDN:AAA34988.1; PID:g172450
R:Pohl, T.M.; Aljinovic, G.
submitted to the Protein Sequence Database, September 1995
A:Reference number: S56798
A:Accession: S56798
A:Molecule type: DNA
A:Residues: 1-399 <TOV>
A:Cross-references: EMBL:Z49301; MID:g1008141; PIDN:CAA89317.1; PID:g1008142; GSPDB:GN00021
C:Genetics:
A:Gene: SGD:RNR2; MIPS:YJL026w
A:Cross-references: SGD:S0003563; MIPS:YJL026w
A:Map position: 10L
C:Function:
A:Description: oxidoreductase; pyrimidine deoxynucleotide metabolism
```

C;Superfamily: ribonucleoside reductase small subunit

C;Keywords: iron; oxidoreductase; pyrimidine deoxynucleotide metabolism
F;145,176,179,239,273,276/Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, His) #54
F;183/Active site: Tyr (stable tyrosyl radical) #status predicted

Alignment Scores:
Pred. No.: 4,78e-98 Length: 399
Score: 1169.50 Matches: 221
Percent Similarity: 81.07% Conservative: 53
Best Local Similarity: 65.38% Mismatches: 58
Query Match: 63.35% Indels: 6
DB: 1 Gaps: 3

US-10-698-228-2 (1-1053) x A26916 (1-399)

```

QY 54 ATCTTCAGACACCAACCAAGTGAATTAAGTCAATGAAGCCACTCTTAAGAAGAG 113
Db 64 ValHisArgHisLysLeuLys-GluMetGlu---LysGluGluProLeuLeuAsnGluAs 82
QY 114 TTCTCGCCGGTTGTTCATCTTCCATCCAGTACCTCGATATTGGAAATGTATAACA 173
Db 82 pLysGluArgThrValLeuPheProIleLysTyrHisGluIleTrpGlnAlaTyrLysAr 102
QY 174 GGCACAGGCTCTCTTCGACAGCAGAGAGTGTGACTTATCAAGAGATCTCCCTCACTG 233
Db 102 gAlaGluAlaSerPheTrpThrAlaGluGluLeuAspLeuSerLysAspIleHisAspTr 122
QY 234 G---AACAGCTTAAAGCAGATGAGAGTACTTCTCATCTCTCATCTTACGCTTTTTCG 290
Db 122 pAsnAsnArgMetAsnGluAsnGluArgPhePheIleSerArgValLeuAlaPheAl 142
QY 291 AGCCAGTGTGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGT 350
Db 142 aAlaSerAspGlyIleValAsnGluAsnLeuValGluAsnPheSerThrGluValGlnIl 162
QY 351 TCCAGAGGCTCGCTGTCTTATGGCTTTTCAAAATTTCTCATCGAAATGTTTCACTCAGAGAT 410
Db 162 eProGluAlaLysSerPheTyrGlyPheGlnIleMetIleGluAsnIleHisSerGluTh 182
QY 411 GTACAGTTGCTGATAGACACTTACATCAGAGTCCCAAGAAAGAGGAATTTTATTAA 470
Db 182 rTyrSerLeuLeuIleAspThrTyrIleLysAspProLysGluSerGluPheLeuPheAs 202
QY 471 TGCATTGAACACATCCCTATGTTTAAGAAAGCAGATTTGGCTTCGCGATCGATAGC 530
Db 202 nAlaIleHisThrIleProGluIleGlyGluLysAlaGluTrpAlaLeuArgTrpIleGl 222
QY 531 AGATAGAAATCTACTTTTGGGAAAGAGTGTGGCTTTTGTCTGTAGAAGAGTGT 590
Db 222 nAspAlaAspAlaLeuPheGlyGluArgLeuValAlaPheAlaSerIleGluGlyValPh 242
QY 591 CTTCTCAGGATCTTTGCTGTATATCTGGCTTAAAGAGAGAGGTCTTATGCCAGACT 650
Db 242 ePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArgGlyMetMetProGlyLe 262
QY 651 CACTTTTTCATGACTCATCAGCAGATGAGGACTTCTACTGTGCTTCTGCTTGCCT 710
Db 262 uThrPheSerAsnGluLeuIleCysArgAspGluGlyLeuHisThrAspPheAlaCysLe 282
QY 711 GATGTTTCAATACTTAGTAAATAAGCTTTCAGAAGAAAGGGTCAGGAGATCATTTGTA 770
Db 282 uLeuPheAlaHisLeuLysAsnLysProAspProAlaIleValGluLysIleValThrGl 302
QY 771 TGCTGTCAAAATGACAGGAGTGTTTTAAACAGAGCCTTGGCAGTTGGCTCATTTGGAAT 830
Db 302 uAlaValGluIleGluGlnArgTyrPheLeuAspAlaLeuProValAlaLeuLeuGlyMe 322
QY 831 GAATTGCATTTTGTGAACACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAAT 890
Db 322 tAsnAlaAspLeuMetAsnGlnTyrValGluPheValAlaAspArgLeuLeuValAlaPh 342
QY 891 TGGATTCTCAAGGTTTTCAGCGAGAAATCTTTTGTGATTTTATGAAACACATTTCTTT 950

```

```

Db 342 eGlyAsnLysLysTyrTyrLysValGluAsnProPheAspPheMetGluAsnIleSerLe 362
QY 951 AGAAGGAAAAACAATTTCTTTAGAAAACGAGTTTCAGAGTATCAGCGCTTTTCAGAGTTAT 1010
Db 362 uAlaGlyLysThrAsnPhePheGluLysArgValSerAspTyrGlnLysAlaGlyValMe 382
QY 1011 GGCAGAAAACACAGAT-----AACGTCTTCCACTTTCGATGCGAGATTTT 1053
Db 382 tSerLysSerThrLysGlnGluAlaGlyAlaPheThrPheAsnGluAspPhe 399

```

RESULT 15

T39992

ribonucleoside-diphosphate reductase small chain - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T39992
R;lyne, M.; Rajandream, M.A.; Barrell, B.G.; Whithead, S.; Chillingworth, T.; Churcher, submitted to the EMBL Data Library, July 1998
A;Reference number: Z21897
A;Accession: T39992
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-391 <LVN>
A;Cross-references: UNIPROT:P36603; EMBL:AL031158; PIDN:CAA20100.1; GSPDB:GN00067; SPDB:
A;Experimental source: strain 972h-; cosmid c25D12
C;Genetics:
A;Gene: SPDB:SPBC25D12.04
A;Map position: 2
C;Superfamily: ribonucleoside reductase small subunit

Alignment Scores:

Pred. No.: 2,54e-97 Length: 391
Score: 1161.50 Matches: 227
Percent Similarity: 78.80% Conservative: 48
Best Local Similarity: 65.04% Mismatches: 61
Query Match: 62.92% Indels: 13
DB: 2 Gaps: 4

US-10-698-228-2 (1-1053) x T39992 (1-391)

```

QY 37 GATCAGGATGAGATCATCTTTCAGAC-----ACCAACGAAAGTGAATAAAG 84
Db 46 GluGluGluLysGlnGluGlyAspTyrTyrLeuGlyLysGluAspGluLeu--- 64
QY 85 TCAATGAAGACGACCTCTTAAGAAAGAGTCTCGCGGTTTGTCACTTTCCATCCAG 144
Db 65 -----AspGluValValLeuArgProAsnProHisArgPheValLeuPheProIleLys 82
QY 145 TACCCTGATATTGGAAATGTATAACAGCAGCAGCTTCTCTTCGACAGCAGAAG 204
Db 83 TyrHisGluIleTrpGlnPheTyrLysLysAlaGluAlaSerPheTrpThrAlaGlu 102
QY 205 GTCGACTTATCAAGGATCTCCCTCACTGG---AACAAAGCTTAAAGCAGATGAGAAGTAC 261
Db 103 IleAspLeuSerLysAspLeuValAspTrpAspAsnLysLeuAsnAlaAspGluArgTyr 122
QY 262 TTCATCTCTCACTTAGCCTTTTTCGACCGCAGTGTGAATTTGAAATGAAATTTG 321
Db 123 PheIleSerThrValLeuAlaTyrPheAlaAlaSerAspGlyIleValAsnGluAsnLeu 142
QY 322 GTCGAGCGCTTAGTCAGGAGTGCAGGTTCAGAGCTCGCTCTTCTATGGCTTCAA 381
Db 143 LeuGluArgPheSerSerGluValGlnIleProGluAlaArgCysValTyrGlyPheGln 162
QY 382 ATTCTCATTCAGAAATGTTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTATCATCAGA 441
Db 163 IleMetIleGluAsnIleHisSerGluThrTyrSerLeuLeuLeuAspThrTyrIleArg 182
QY 442 GATCCCAAGAAAAGGGAATTTTATTATTAATGCAATTAAGAACCATGTCCTATGTTAAGAAA 501
Db 183 GluProLysGluLysGlnArgHisPheAspAlaIleLeuThrMetGlySerIleLysAla 202
QY 502 AAAGCAGATTGGCGCTTGGCATCGATAGCAGATAAAATCTACTTTTGGGGAAGAGTG 561

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 30, 2005, 03:28:20 ; Search time 76 Seconds
(without alignments)

14189.984 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1846

Sequence: 1 atgggcgaccggaaggcc.....tcacctggatgcagatttt 1053

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10698228/runat_26102005_100608_3981/app_query.fasta.1.2446
-DB=UniProt_03 -QMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10698228@cgn_1_1_149@runat_26102005_100608_3981 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	1821	98.6	351	2	Q7LG56 homo sapien
2	1817	98.4	351	2	Q9NUW3
3	1725	93.4	351	2	Q6PEE3
4	1534	83.1	297	2	Q6AI41
5	1514	82.0	299	2	Q75PY9
6	1474	79.8	285	2	Q8VEE3
7	1454	78.8	389	1	RIR2_HUMAN
8	1443	78.2	390	1	RIR2_MOUSE
9	1439	78.0	386	2	Q6EEF1
10	1431	77.5	386	2	Q801Q4
11	1427	77.3	386	2	Q8AVT2
12	1425.5	77.2	386	1	RIR2_BRARE
13	1421	77.0	386	1	RIR2_MESAU
14	1420.5	77.0	386	2	Q6DI44
15	1399.5	75.8	378	2	Q6P876
16	1385	75.0	406	2	Q6IP47

17	1382.5	74.9	388	2	Q7PF28
18	1382.5	74.9	426	2	Q7QIF4
19	1370	74.2	349	2	Q7ZYW0
20	1358	73.6	397	2	Q95VP8
21	1345.5	73.1	399	2	Q27124
22	1348	73.0	393	1	RIR2_DROME
23	1335	72.3	398	2	Q9XYN8
24	1331	72.1	384	1	RIR2_SPIPO
25	1284	69.6	403	2	Q6CFU6
26	1280	69.3	319	2	O57175
27	1280	69.3	319	2	Q76Q46
28	1280	69.3	319	2	Q89087
29	1279	69.3	319	1	RIR2_VACCV
30	1279	69.3	319	2	Q76ZX1
31	1277	69.2	319	2	Q6RZQ9
32	1276	69.1	319	1	RIR2_VACCC
33	1276	69.1	319	2	Q8JLH6
34	1276	69.1	319	2	Q89559
35	1274	69.0	333	2	Q8QN14
36	1273	69.0	319	1	RIR2_VACCP
37	1271	68.9	319	2	Q8V544
38	1270	68.8	319	2	P87632
39	1268	68.7	319	2	Q8V2X8
40	1268	68.7	319	2	Q77525
41	1267	68.6	319	2	Q9JFE9
42	1266	68.6	319	1	RIR2_VARV
43	1264.5	68.5	410	1	RIR2_NEUCR
44	1260	68.3	333	2	Q9PXK4
45	1258.5	68.2	348	2	Q9YMK7

ALIGNMENTS

RESULT 1

Q7LG56 PRELIMINARY; PRT; 351 AA.
ID Q7LG56 Q9NPDS; Q9NTD8;
AC Q7LG56 Q9NPDS; Q9NTD8;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)
DE Ribonucleotide reductase (Hypothetical protein DKFZp761E1312) (P53-
inducible ribonucleotide reductase small subunit 2).
OS Homo sapiens (Human).
GN Name=p53R2; Synonyms=DKFZp761E1312;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20179179; PubMed=10716435; DOI=10.1038/35003506;
RA Tanaka H., Arakawa H., Yamaguchi T., Shiraishi K., Fukuda S.,
RA Matsui K., Takei Y., Nakamura Y.,
RT "A ribonucleotide reductase gene involved in a p53-dependent cell-
cycle checkpoint for DNA damage.";
RL Nature 404:42-49(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Amygdala;
RG The German cDNA Consortium;
RA Ansoerge W., Krieger S., Regiert T., Rittmuller C., Schwager B.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Ugai H., Yokoyama K.K.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB036524; BAA92493.1; JOINED.
DR EMBL; AB036525; BAA92493.1; JOINED.
DR EMBL; AB036526; BAA92493.1; JOINED.
DR EMBL; AB036527; BAA92493.1; JOINED.
DR EMBL; AB036528; BAA92493.1; JOINED.
DR EMBL; AB036529; BAA92493.1; JOINED.

Q7pf28 anopheles g
Q7qif4 anopheles g
Q7zyw0 brachydanio
Q95vp8 aedes aegypt
Q27124 urechis cau
P48592 drosophila
Q9xyN8 aedes albop
P07201 spissula sol
O6cfu6 yarrowia li
O57175 vaccinia vi
Q76q46 variola min
Q89087 variola vir
P1158 vaccinia vi
Q76zx1 vaccinia vi
Q6rzo9 rabbitpox v
P20493 vaccinia vi
Q8jln6 ectromelia
Q89559 variola vir
Q8qn14 cowpox viru
P29883 vaccinia vi
Q8v544 monkeypox v
P87632 cowpox viru
Q8v2x8 camelpox vi
Q77525 camelpox vi
Q9jfe9 vaccinia vi
P33799 variola vir
Q9cl67 neurospora
Q9pxr4 variola vir
Q9ymk7 lymantria d

DR GO; GO:0004748; F:ribonucleoside-diphosphate reductase activity; IEA.
DR GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonucleotide reductase.
DR Pfam; PF00268; Ribonucleotide reductase.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
SQ SEQUENCE 351 AA; 40704 MW; 6D009B2D59E9A323 CRC64;

Alignment Scores:
Pred. No.: 9.34e-151 Length: 351
Score: 1817.00 Matches: 350
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.72% Mismatches: 0
Query Match: 98.43% Indels: 0
DB: 2 Gaps: 0

US-10-698-228-2 (1-1053) x Q9NUW3 (1-351)

Qy 1 ATGGGCGACCGGAAAGCCGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Qy 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTGC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
Qy 121 CGGTTTCGTCATCTTCCAAATCCAGTACCTGATATTTGGAAAATCTATAACACGACAG 180
Db 41 ArgPheValIlePheProIleGlnTyProAspIleTrpLysMetTyLysGlnAlaGln 60
Qy 181 GCTTCTCTTCGACAGCAGAGAGGTGCATCTATCAAGAGGATCCCTCACTGGAACAAG 240
Db 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Qy 241 CTTAAAGCAGATGAAGTACTTCTCATCTCTCACTTAGCCTTTTTCGACCCAGTGTAT 300
Db 81 LeuLysAlaAspGluLysTyPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Qy 301 GGAATTCGAATGAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Qy 361 CGCTGTTCTATGCTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTySerLeu 140
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTCAA 480
Db 141 LeuIleAspThrTyIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Qy 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCGCTTGCATGATGATGAGATAGAAA 540
Db 161 ThrMetProTyValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTCTGCTGTGTAGAGAGTTCCTCTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Qy 601 TCTTTTCTGCTATATTCTGGCTAAAGAAGAGGTCTTATGCCAGGACTCACATTTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Qy 661 AATGAACATCATCAGCAGATGAAGGATTCACGTGTCATCTTGTCTGCTGATGTTCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Qy 721 TACTTAGTAATAGCTTTCAGAAAGAGGTTCAGGAGATCATTTGTTGCTGCTGCTCAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCCTTGCAGTTGCGCTCATTTGGAATGATTCATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyValAsnCysIle 280

Qy 841 TTGATGAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 900
Db 281 LeuMetLysGlnTyIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Qy 901 AAGTTTTTCAGGCGAGAAAATCCTTTGATTATTTAGAAAACATTTCTTTAGAGGAAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Qy 961 ACAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGCTTATGCGAGAAC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyGlnArgPheAlaValMetAlaGluThr 340
Qy 1021 ACAGATAACGCTTTCACCTTGGATGCGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 3
Q6PEE3 PRELIMINARY; PRT; 351 AA.
AC Q6PEE3;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Ribonucleotide reductase M2 B (TP53 inducible).
GN Name=Rxm2b;
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC058103; AAHS8103.1; "
DR GO; GO:0004748; P:ribonucleoside-diphosphate reductase activity; IEA.
DR GO; GO:0009263; P:deoxyribonucleotide biosynthesis; IEA.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonucleotide reductase.
DR Pfam; PF00268; Ribonucleotide reductase.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
SQ SEQUENCE 351 AA; 40803 MW; 4E1259233C9CC8A9 CRC64;

Alignment Scores:
Pred. No.: 1.13e-142 Length: 351
Score: 1725.00 Matches: 328
Percent Similarity: 97.44% Conservative: 14
Best Local Similarity: 93.45% Mismatches: 9

Query Match:	93.45%	Indels:	0
DB:	2	Gaps:	0
US-10-698-228-2 (1-1053) x Q6PBE3 (1-351)			
QY	1	ATGGGACCGCGGAAAGCCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA	60
Db	1	MetGlyAspProGluArgProGluAlaAlaArgProGluGlyGluGlnLeuCysSer	20
QY	61	GACACCAACGAAGTGAATAAGTCAAAATCAAGAGCCACTCCCTAAAGAAAGAGTTCTCGC	120
Db	21	GluThrGluGlnValValArgSerAsnGluGluProLeuLeuArgLysSerSerArg	40
QY	121	CGGTTTGTCTATTTCCAAATCCAGTACCTCGATATTTGGAAAATGTATAAAGCCACAG	180
Db	41	ArgPheValIlePheProIleGlnTyrProAspIleTrpArgMetTyrLysGlnAlaGln	60
QY	181	GCTTCCTCTGGCAGCAGAGAGGTGCACTATCAAGGATCTCCCTCACTGAACAG	240
Db	61	AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys	80
QY	241	CTTAAAGCAGATGAGAGTAGTCTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTAT	300
Db	81	LeuLysSerAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp	100
QY	301	GGAATTGTAATAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGCT	360
Db	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla	120
QY	361	CGCTGTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTG	420
Db	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140
QY	421	CTGATGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA	480
Db	141	LeuIleAspThrTyrIleArgAspProLysArgGluPheLeuPheAsnAlaIleGlu	160
QY	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCGCTTGGCATGGATGAGATAGAAAA	540
Db	161	ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys	180
QY	541	TCTACTTTGGGAAAGAGTGGTGGCTTGTCTGCTGTAGAGAGTTTCTTCTCAGGA	600
Db	181	SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyIlePhePheSerGly	200
QY	601	TCTTTTGTCTATATTCGTGCTAAAGAGAGGTCTTATGCCAGGACTCACCTTTTCC	660
Db	201	SerPheAlaAlaIlePheTrpLeuLysArgGlyLeuMetProGlyLeuThrPheSer	220
QY	661	AATGAATCTATCAGCAGATGAAGGACTTCACCTGTGACTTTGCTGCCTGATGTTCCAA	720
Db	221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln	240
QY	721	TACTTAGTAATAAGCCTTCAGAGAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA	780
Db	241	TyrLeuValAsnLysProSerGluAspArgValArgGluIleAlaAspAlaValGln	260
QY	781	ATTGACGAGGTTTAAACAGAGCCTGCCAGTTCGCTCATTTGGAATCAATTGCATT	840
Db	261	IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysVal	280
QY	841	TTGATGAAACAGTACATTGAGTTTGTAGCTGCACAGATTACTTGTGGAACTTGGATTCTCA	900
Db	281	LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuGlyGluLeuGlyPheSer	300
QY	901	AAGGTTTTTCAGCAGAAAAATCTTTTGAATTTATGAAAAACATTTCTTTAGAGGAAAA	960
Db	301	LysIlePheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys	320
QY	961	ACAAATTTCTTCAGAAACAGTTTCAGAGTATCAGGGTTTTCAGTTATGGCAGAACCC	1020
Db	321	ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr	340
RESULT 4			
QY	1021	ACAGATAACGCTTCTCACCTTGGATCAGATTTT 1053	
Db	341	ThrAspAsnValPheThrLeuAspAlaAspPhe 351	
PRELIMINARY; PRT; 297 AA.			
ID	Q6AI41		
AC	Q6AI41		
DT	25-OCT-2004 (TRENBLrel. 28, Created)		
DT	25-OCT-2004 (TRENBLrel. 28, Last sequence update)		
DT	25-OCT-2004 (TRENBLrel. 28, Last annotation update)		
DE	Hypothetical protein DKFP686M05248.		
GN	Name=DKFP686M05248;		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RC	SEQUENCE FROM N.A.		
RP	TISSUE=Salivary gland;		
RG	The German cDNA Consortium;		
RA	Ottewaelder B., Obermaier B., Deutschenbaur S., Schaipp A.,		
RA	Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;		
RL	Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; CR627376; CAH10473.1; -		
DR	GO; GO:0004748; F:ribonucleoside-diphosphate reductase activity; IEA.		
DR	GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.		
DR	InterPro; IPR009078; Ferritin/RR like.		
DR	InterPro; IPR000358; Ribonucleotidse.		
DR	Pfam; PF0268; Ribonuc_red_sm; 1.		
DR	PROSITE; PS00368; RIBORED_SMALL; 1.		
KW	Hypothetical protein.		
SQ	SEQUENCE 297 AA; 34498 MW; 2B5522F2CC158A02 CRC64;		
Alignment Scores:			
Align. No.:	6.61e-126	Length:	297
Score:	1534.00	Matches:	296
Percent Similarity:	99.66%	Conservative:	0
Best Local Similarity:	99.66%	Mismatches:	1
Query Match:	83.10%	Indels:	0
DB:	2	Gaps:	0
US-10-698-228-2 (1-1053) x Q6AI41 (1-297)			
QY	163	ATGTATAAACAGGACAGGCTTCTCTGGCAGCAGAGAGTCTGACTTATCAAAGGAT	222
Db	1	MetTyrLysGlnAlaGlnAlaSerPheTrpThrAlaGluValAspLeuSerLysAsp	20
QY	223	CTCCCTCACTGGAACAAGCTTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCC	282
Db	21	LeuProHisTrpAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAla	40
QY	283	TTTTTTCAGGAGTGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAG	342
Db	41	PhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGlu	60
QY	343	GTGCAGGTTCCAGAGGCTCGCTGTTTCTTATGCTTTCAAATTTCTCATCCAGAAATGTTAC	402
Db	61	ValGlnValProGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHis	80
QY	403	TCAGAGATGTACAGTTTCTGATAGACATTTACATCAGAGATCCCCAGAAAAAGGGAATTT	462
Db	81	SerGluMetTyrSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPhe	100
QY	463	TTATTTAATGCAATTGAAACCATGCCCTATGTTAAGAAAAAGCAGATTGGCGCTTGGCA	522
Db	101	LeuPheAsnAlaIleGluThrMetProTyrValLysLysAlaAspTrpAlaLeuArg	120
QY	523	TGGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCTTCTCTCTCTGTAGAA	582
Db	121	TrpIleAlaAspArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGlu	140
QY	583	GGAGTTTTTCTCTCAGGATCTTTTGTGCTATATTCTGCTAAAGAGAGAGGTCTTATG	642

Db	141	GlyValPhePheSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMet	160
QY	643	CCAGACTCACCTTTTCCAAATGAACTCATCAGCAGAGATGAAGGACTTCACCTGTGACTTTT	702
Db	161	ProGlyLeuThrPheSerAsnGluLeuIleSerArgGluGlyLeuHisCysAspPhe	180
QY	703	GCTGCCTGATGTTCCAACTACTAGTAATAATAGCCTTCAGAAAGAGGTCAGGAGATC	762
Db	181	AlaCysLeuMetPheGlnTrpLeuValAsnLysProGluGluArgValArgGluIle	200
QY	763	ATTGTTGATGCTCAAAATTCAGCAGAGAGTTTAAACAGAGCCCTGCGAGTTGGCCTC	822
Db	201	IleValAspAlaValLysIleGluGlnPheLeuThrGluAlaLeuProValGlyLeu	220
QY	823	ATTGGAATGAATGCTATTTGATGAACAGTACATTTAGTTGTAGCTACACAGATTACTT	882
Db	221	IleGlyMetAsnCysIleLeuMetLysGlnTrpIleGluPheValAlaAspArgLeuLeu	240
QY	883	GTGGAACCTGGATTCTCAAGGTTTTTTCAGGCAGAGAAATCCTTTTGATTTTATGAAAC	942
Db	241	ValGluLeuGlyPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsn	260
QY	943	ATTCTTTTAGAAGGAAACAAATTTCTTTTGAGAACGAGTTTCAGAGTATCAGCGTTT	1002
Db	261	IleSerLeuGluGlyLysThrAsnPhePheGluLysArgValSerGluTrpGlnArgPhe	280
QY	1003	GCAGTTATGGCAGAAACACACAGATAACGCTCTTCACCTTGGATGCAGATTTT	1053
Db	281	AlaValMetAlaGluThrThrAspAsnValPheThrLeuAspAlaAspPhe	297
RESULT 5			
ID	Q75PY9	PRELIMINARY;	PRT; 299 AA.
AC	Q75PY9;		
DT	05-JUL-2004 (TReMBLrel. 27, Created)		
DT	05-JUL-2004 (TReMBLrel. 27, Last sequence update)		
DE	P53-inducible ribonucleotide reductase small subunit 2 long form.		
GN	Name=p53R2;		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Ugai H., Yokoyama K.K.;		
RL	Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AB163437; BAD11774.1; -		
DR	GO; GO:0004748; P:ribonucleoside-diphosphate reductase activity; IEA.		
DR	GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.		
DR	InterPro; IPR009078; Ferritin/AR-like.		
DR	InterPro; IPR000358; Ribonucleotide reductase.		
DR	Pfam; PF00268; Ribonucleotide reductase.		
DR	PROSITE; PS00368; RIBORED SMALL; 1.		
SQ	SEQUENCE 299 AA; 34528 MW; 4705C44389EB689B CRC64;		
Alignment Scores:			
Pred. No.:	3,78e-124	Length:	299
Score:	1514.00	Matches:	299
Percent Similarity:	85.19%	Conservative:	0
Best Local Similarity:	85.19%	Mismatches:	0
Query Match:	82.02%	Indels:	52
DB:	2	Gaps:	1
US-10-698-228-2 (1-1053) x Q75PY9 (1-299)			
QY	1	ATGGCGAGCCCGAAGCCCGAGCGCGCGCTGATCAGGATCAGGATCATCTTCA	60
Db	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAsp	15
QY	61	GACACCAACGAAAGTGAATAAAGTCAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTTCGC	120

DE Similar to ribonucleotide reductase M2 polypeptide (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC042468; AAH42468.1; -.
DR HSP; P11157; IYXSM.
DR GO; GO:0004748; F:ribonucleoside-diphosphate reductase activity; IEA.
DR GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonuc redctase.
DR Pfam; PF00268; Ribonuc red sm-1.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
FT NON TER 1
SQ SEQUENCE 285'AA; 33075 MW; C4E7D536479BC15B CRC64;

Alignment Scores:
Pred No.: 1-23e-120 Length: 285
Score: 1474.00 Matches: 285
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 79.85% Indels: 0
DB: 2 Gaps: 0

US-10-698-228-2 (1-1053) x Q86VE3 (1-285)
QY 199 GAAGAGCTGCAGCTTATCAAGAGTCTCCCTCAGTGAACAGCTTAAAGCAGATGAGAAG 258
Db 1 GluGluValAspLeuSerLysAspLeuProHisThrAsnLysLeuLysAlaAspGluLys 20
QY 259 TACTTCATCTCTCATCTTAGCTTACCTTTTTCAGCCAGTGTGAAATGTAATGAAAT 318
Db 21 TyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsn 40
QY 319 TTGGTGAGCGCTTTAGTTCAGGAGTGCAGGTCCTCAGAGGCTCGCTGTTCTATGGCTTT 378
Db 41 LeuValGluArgPheSerGlnGluValGlnValProGluAlaAlaArgcysPheTyrGlyPhe 60
QY 379 CAAATTCCTCATCGAAGTGTTCCTCAGAGATGTACAGATTTGCTGATGACACTTACATC 438
Db 61 GlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeuLeuIleAspThrTyrIle 80
QY 439 AGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAACCATGCTCTATGTTAAG 498
Db 81 ArgAspProLysLysArgGluPheLeuPheAsnAlaIleGluThrMetProTyrValLys 100
QY 499 AAAAAACAGATTGGGCTTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 558
Db 101 LysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLysSerThrPheGlyGluArg 120
QY 559 GTGGTGCCCTTTGCTGCTGTAGAGGAGTTCCTCTCAGGATCTTTGCTGCTATATTC 618
Db 121 ValValAlaPheAlaAlaValGluGluValPhePheSerGlySerPheAlaAlaIlePhe 140
QY 619 TGGCTAAAGAGAGAGTCTTATGCCAGGACTCATCTTTTCCATGAACTCATCAGCAGA 678
Db 141 TrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArg 160
QY 679 GATCAAGAGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 738
Db 161 AspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGlnTyrLeuValAsnLysPro 180
QY 739 TCAGAGAAAGGGTCAGGAGATCATTTGTCATGCTGCTCAAAATTCAGCAGGAGTTTGA 798
Db 181 SerGluGluArgValArgGluIleIleValAspAlaValLysIleGluGlnGluPheLeu 200
QY 799 ACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATTCATTTGATGAAGACAGTACATT 858
Db 1

Db 201 ThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIleLeuMetLysGlnTyrIle 220
QY 859 GAGTTTGTAGCTGACAGATTACTTCTGGAACTTGGATTCTCAAGGCTTTTTCAGGCAGAA 918
Db 221 GluPheValAlaAspArgLeuLeuValGluLeuGlyPheSerLysValPheGlnAlaGlu 240
QY 919 AATCTTTTGTATTTATGAAACATTTCTTTAGAGGAAAAACAATTTCTTTTGAGAAA 978
Db 241 AsnProPheAspPheMetGluAsnIleSerLeuGluGlyLysThrAsnPheGluLys 260
QY 979 CGAGTTTCAGAGTATCAGCGTTTTCAGTATGATGCGAAGAACACAGATACCTTTCCAC 1038
Db 261 ArgValSerGluTyrGlnArgPheAlaValMetAlaGluThrThrAspAsnValPheThr 280
QY 1039 TTGGATGCGAGATTTT 1053
Db 281 LeuAspAlaAspPhe 285

RESULT 7
RIR2 HUMAN
ID_RIR2 HUMAN STANDARD; PRT; 389 AA.
AC P31350;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Ribonucleoside-diphosphate reductase M2 chain (EC 1.17.4.1)
DE (Ribonucleotide reductase small chain).
GN Name=RRM2; Synonyms=RR2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast carcinoma;
RX MEDLINE=92329977; PubMed=1627826;
RA Pavloff N., Rivard D., Masson S., Shen S.-H., Mes-Masson A.-M.;
RT "Sequence analysis of the large and small subunits of human
ribonucleotide reductase.";
RL DNA Seq. 2:227-234(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21974633; PubMed=11978970;
RA Zhou B., Yen Y.;
RT "Characterization of the human ribonucleotide reductase M2 subunit
gene; genomic structure and promoter analyses.";
RL Cytogenet. Cell Genet. 95:52-59(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle, and Skin;
RX MEDLINE=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villallon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Provides the precursors necessary for DNA synthesis.
CC -!- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate +

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87089677; PubMed=3025593;
RA Thelander L., Berg P.;
RT "Isolation and characterization of expressible cDNA clones encoding
RL the M1 and M2 subunits of mouse ribonucleotide reductase.";
RN Mol. Cell. Biol. 6:3433-3442(1986).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90060004; PubMed=2684652;
RA Thelander M., Thelander L.;
RT "Molecular cloning and expression of the functional gene encoding the
RL M2 subunit of mouse ribonucleotide reductase: a new dominant marker
RN gene.";
RN EMBO J. 8:2475-2479(1989).
[3]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 65-352.
RX MEDLINE=97030707; PubMed=8876648; DOI=10.1006/jmbi.1996.0546;
RA Kauppi B., Nielsen B.B., Ramaswamy S., Larsen I.K., Thelander M.,
RA Thelander L., Eklund H.;
RT "The three-dimensional structure of mammalian ribonucleotide reductase
RL protein R2 reveals a more-accessible iron-radical site than
RT Escherichia coli R2.";
RL J. Mol. Biol. 262:706-720(1996).
[4]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 65-352.
RX MEDLINE=22206532; PubMed=12087093; DOI=10.1074/jbc.M203358200;
RA Strand K.R., Karlsson S., Andersson K.K.;
RT "Cobalt substitution of mouse R2 ribonucleotide reductase as a model
RL for the reactive diferrous state. Spectroscopic and structural
RT evidence for a ferromagnetically coupled dinuclear cobalt cluster.";
RL J. Biol. Chem. 277:34229-34238(2002).
[5]
RP STRUCTURE BY NMR OF 384-390.
RX MEDLINE=96069780; PubMed=7583667;
RA Fisher A.L., Laub P.B., Cooperman B.S.;
RT "NMR structure of an inhibitory R2 C-terminal peptide bound to mouse
RL ribonucleotide reductase R1 subunit.";
RL Nat. Struct. Biol. 2:951-955(1995).
CC -!- FUNCTION: Provides the precursors necessary for DNA synthesis.
CC -!- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate +
CC thioredoxin disulfide + H(2)O = ribonucleoside diphosphate +
CC thioredoxin.
CC -!- COPACTOR: Binds 2 iron ions per subunit.
CC -!- PATHWAY: DNA replication pathway; first step.
CC -!- SUBUNIT: Heterodimer of a large (M1) and a small chain (M2).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- MISCELLANEOUS: Two distinct regulatory sites have been defined:
CC the specificity site, which controls substrate specificity, and
CC the activity site, which regulates overall catalytic activity. A
CC substrate-binding catalytic site, located on M1, is formed only in
CC the presence of the second subunit M2.
CC -!- SIMILARITY: Belongs to the ribonucleoside diphosphate reductase
CC small chain family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M14223; AAA40062.1; -;
DR EMBL; X15666; CAA33707.1; -;
DR PIR; S06735; S06735.
DR PDB; 1AFT; NMR; @=383-390.
DR PDB; 1H0N; X-ray; A=1-390.
DR PDB; 1H0O; X-ray; A=1-390.
DR PDB; 1XSM; X-ray; @=1-390.
DR MGI; 98181; Rrm2.
DR GO; GO:0005515; F:protein binding; IPI.

DR GO; GO:0009262; P:deoxyribonucleotide metabolism; IDA.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR003358; Ribonucl_redctse.
DR Pfam; PF00268; Ribonuc_red_sm_1.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
KW 3D-structure; DNA replication; Iron; Metal-binding; Oxidoreductase.
FT METAL 139 139 Iron 1.
FT METAL 170 170 Iron 1 and 2.
FT METAL 173 173 Iron 1.
FT METAL 233 233 Iron 2.
FT METAL 267 267 Iron 1 and 2.
FT METAL 270 270 Iron 2.
FT ACT SITE 177 177 By similarity.
FT HELIX 68 70
FT TURN 72 74
FT HELIX 89 99
FT TURN 100 101
FT HELIX 105 107
FT TURN 112 113
FT HELIX 114 118
FT TURN 119 119
FT HELIX 122 148
FT TURN 149 149
FT HELIX 150 153
FT HELIX 157 184
FT HELIX 188 195
FT TURN 196 196
FT HELIX 197 200
FT HELIX 202 204
FT HELIX 205 216
FT TURN 218 219
FT HELIX 222 235
FT TURN 236 237
FT HELIX 238 249
FT TURN 250 251
FT HELIX 254 279
FT TURN 280 280
FT HELIX 287 306
FT TURN 307 308
FT HELIX 311 314
FT TURN 315 315
FT HELIX 318 335
FT TURN 336 337
FT HELIX 349 351
SQ SEQUENCE 390 AA; 45095 MW; AC7ACC4FAF8A4A2F CRC64;

Alignment Scores:
Pred. No.: 6.77e-118 Length: 390
Score: 1443.00 Matches: 269
Percent Similarity: 90.62% Conservative: 40
Best Local Similarity: 78.89% Mismatches: 26
Query Match: 78.17% Indels: 6
DB: 1 Gaps: 1

US-10-698-228-2 (1-1053) x RIR2_MOUSE (1-390)
QY 49 AGATCATCTTCAGACACCAACAGTAAGTCAATTAAGTCA-----AAT 90
Db 50 ArgilePheGlnAspSerAlaGluLeuSerLysAlaProThrAsnProSerValGlu 69
QY 91 GAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGGTTGTCTATCTTTTCCAATCCAGTACCT 150
Db 70 AspGluProLeuLeuArgGluAsnProArgPheValValPheProIleGluTyrHis 89
QY 151 GATATTGGAAATGTATAACAGGCACAGGCTTCCTTCTGGACAGCAGAGAGTCCGAC 210
Db 90 AspIleTrpGlnMetTyrLysLysAlaGluAspPheTrpThrAlaGluValAsp 109
QY 211 TTATCAAGAGGATCTCCCTCAGTGAACAGCTTAAAGCAGATGAGAAGTACTTCTCTCT 270
Db 110 LeuSerLysAspIleGlnHisTrpGluAlaLeuLysProAspGluArgHisPheIleSer 129
QY 271 CACATCTTAGCCTTTTGTGACGCCAGTGAATGGAATTTGAAATTTTGTGTGGAGCC 330

Db	130	HisValLeuLeuAlaPhePheAlaAAserAspGlyIleValAsnGluAsnLeuValGluArg	149
Qy	331	TTTAgTcAGAGGtGcAGGtTCCAGAGGcTcGcTgTtTCTATGGcTtTCAAAATtCTcATC	390
Db	150	PheSerGlnGluValGlnValThrGluAlaArgCysPheTyrGlyPheIleAlaMet	169
Qy	391	GAGAAcTGTcACTcAGAGATgTACAGATtTGTGTAGAcACTTAcATcAGAGATcCCcAAG	450
Db	170	GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys	189
Qy	451	AAAAGGAAATTTTATTTAAATGCcAATTCAAACcATGCcCTATgTTTAAAGAAAAGcAGAT	510
Db	190	GluArgGluTyrLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAsp	209
Qy	511	TGGcCTTGGcATgATGAcAGATAGAAAATcTACTTTTGGGAAAGAGTGGTgCCcTTT	570
Db	210	TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyLysValValAlaPhe	229
Qy	571	GCTGcTGTAGAAAGAGTtTTCTTCTcAGAGTcTTTtTGTCTATATtTGTGCTAAAGAAAG	630
Db	230	AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys	249
Qy	631	AGAGGcTCTTAGCCAGGAcTcACTTTTTCcAAATGAAcTcATcAGcAGAGATGAAGAcTT	690
Db	250	ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu	269
Qy	691	CACtGTGAcTtTGTCTGcCTGATgTTCcAAATAcTtTAGTAAATAGcCTTcAGAAAGG	750
Db	270	HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProAlaGluIleArg	289
Qy	751	GTcAGGAGATcATgTGTGCTGCTcAAATtTGAcAGGAGTtTTTAAcAGAGAcCTTG	810
Db	290	ValArgGluIleIleThrAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu	309
Qy	811	CCAGTtGGcCTcATTGGAAATGAATtTGcATtTTTGATGAACAGTAcATTGAGTtTGAcT	870
Db	310	ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla	329
Qy	871	GAcAGATtACTTGTGGAAcTcTGGAAATcTCAAGGtTTTTCAGGcAGAAAAATcCTTTTGAT	930
Db	330	AspArgLeuMetLeuGluLeuGlyPheAsnLysIlePheArgValGluAsnProPheAsp	349
Qy	931	TTTATGAACAcATtCTCTTAGAGGAAAAcAAATtCTTTGAGAAcAGATtTCAGAG	990
Db	350	PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu	369
Qy	991	TATcAGGcTtTGTGcAGTtATGGcAGAAACcAGATAcCTTCTcACCTTGGATGcAGAT	1050
Db	370	TyrGlnArgMetGlyValMetSerAsnSerThrGluAsnSerPheThrLeuAspAlaAsp	389
Qy	1051	TTT 1053	
Db	390	Phe 390	

RESULT 9

ID	Q68EP1	PRELIMINARY;	PRT; 386 AA.
AC	Q68EP1		
AD	Q68EP1;		
DT	25-OCT-2004	(TrEMBLrel. 28, Created)	
DT	25-OCT-2004	(TrEMBUpdrel. 28, Last sequence update)	
DT	25-OCT-2004	(TrEMBUpdrel. 28, Last annotation update)	
DE	Hypothetical protein.		
OS	Xenopus tropicalis (Western clawed frog) (<i>Silurana tropicalis</i>).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;		
OC	Xenopodinae; Xenopus.		
NCBI_TaxID=8364;			
RN	[1]		
RC	SEQUENCE FROM N.A.		
RP	TISSUE=Embryo;		
RX	PubMed=12477932; DOI=10.1073/pnas.242603899;		
RA	Strausberg R.L., Feinsold E.A., Grouse L.H., Derge J.G.,		

Db 211 IleSerAspLysGlnAlaThrTyrGlyGluArgValValAlaPheAlaValGluGly 230
QY 586 GTTTCTCTCAGGATCTTTGCTGCTATATCTGCTAAAGAGAGAGTCTTATGCCA 645
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 231 IlePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArgGlyLeuMetPro 250
QY 646 GGACTCATTCTTCCAAATGACATCATCAGAGATGAAGGACTTCACCTGTGACTTTGCT 705
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 251 GlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAla 270
QY 706 TGCCTGATGTTCCATCTAGTAAATAGCTTCAGAGAAAGGCTCAGGAGATCATTT 765
Db 271 CysLeuMetPheArgHisLeuValHisLysProSerGluGluArgValValGlnLeuIle 290
QY 766 GTTGATGCTGCTCAAAATTCAGCAGGAGTTTAAACAGAAAGCCTTGCAGTTGCGCTCATTT 825
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 291 ThrAsnAlaValGlnIleGluGlnGluPheLeuThrGluAlaLeuProValAsnLeuIle 310
QY 826 GGAATGAATTCATTTTGTATGAACAGTACATGAGTTTGTAGCTGACAGATTTACTTTGTG 885
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 311 GlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuLeu 330
QY 886 GAACCTGATCTCAAGTCTTTCAGGCAGCAAAATCTTTGATTTTATGAAACATTT 945
Db 331 GluLeuGlyPheAsnLysValPheLysAlaSerAsnProPheAspPheMetGluAsnIle 350
QY 946 TCTTTAGAGGAGAAACAAATTTCTTTGAGAACAGATTTTCAGAGTATCAGCGTTTGTGCA 1005
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 351 SerLeuGluGlyLysThrAsnPheGluLysArgValGlyGlnLysMetGly 370
QY 1006 GTTATGCGCAGAACACACATCACTCTTCCACTTGTGATGTCAGATTTT 1053
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 371 ValMetSerLysProLysAspLeuThrPheThrLeuAspAlaAspPhe 386
QY Q8AVY2 PRELIMINARY; PRT; 386 AA.
AC Q8AVY2; ID Q8AVY2; AC Q8AVY2; AC Q8AVY2; AC Q8AVY2; AC Q8AVY2; AC Q8AVY2; AC Q8AVY2;
DT 01-WAR-2003 (TREMELrel. 23, Created)
DT 01-WAR-2003 (TREMELrel. 23, Last sequence update)
DE 01-WAR-2004 (TREMELrel. 26, Last annotation update)
DE MGCS2676 protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.

RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.,
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative.";
RL Dev. Dyn. 225:384-391 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX Klein S., Strausberg R.,
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041209; AAH41209.1; -;
DR HSSP; P1157; 1XSM.
DR GO; GO:0004748; P:Ribonucleoside-diphosphate reductase activity; IEA.
DR GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.
DR InterPro; IPR009078; Ferritin/RR-like.
DR InterPro; IPR000358; Ribonucleotide.
DR Pfam; PF00268; Ribonuc red sm; 1.
DR PROSITE; PS00368; RIBORED SMALL; 1.
SQ SEQUENCE 386 AA; 44635 MW; 4635FCC9BB18EA32 CRC64;
Alignment Scores:
Pred. No.: 1.72e-116 Length: 386
Score: 1427.00 Matches: 266
Percent Similarity: 89.88% Conservative: 36
Best Local Similarity: 79.17% Mismatches: 30
Query Match: 77.30% Indels: 4
DB: 2 Gaps: 1
US-10-698-228-2 (1-1053) x Q8AVY2 (1-386)
QY 58 TCAGACACCAACGAAAGTGAATAAGTCAAT-----GAAGAGCCACTCCTA 105
Db 51 ThrGluThrProLysSerLysAlaProLysAsnProArgLeuGluAspGluProLeuLeu 70
QY 106 AGAAGAGTTCTCGCGGTTGTCTCTTCCATCCAGTACCTCGATATTTGGAAATG 165
::: ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 71 LysAspAsnProHisArgPheValIlePheProIleGlnTyrHisAspIleTrpGlnMet 90
QY 166 TATAAAGCAGCAGGCTTCTTCTGGACGACGAGAGGTCGACTTATCAAGAGATCTC 225
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 91 TyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeu 110
QY 226 CCTCACTGGACAGCTTAAGCAGATGAGAGTACTCTCATCTCATCTTACCTTT 285
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 111 GlnHisTrpGluSerLysLysGluGluLysTyrPheIleSerHisValLeuAlaPhe 130
QY 286 TTTGCAGCCAGTGTGGAATTTGAAATTTGGTGGAGCGCTTTAGTCAGAGGTG 345
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 131 PheAlaLysAspGlyIleValAsnGluAsnLeuValGluArgPheSerLysGluVal 150
QY 346 CAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGAATGTTCACTCA 405
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 151 GlnValThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMetGluAsnIleHisSer 170
QY 406 GAGATGTACAGTTTGTGTAGTACACTTACATCAGAGATCCCAAGAAAAGGAATTTTAA 465
Db 171 GluMetTyrSerLeuLeuIleAspThrTyrValLysAspProLysGluArgGluTyrLeu 190
QY 466 TTTAATGCAATTCAGAACCATGCGCTATGTTAAGAAAAAGCAGATTCGGCTTCGGATGG 525
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 191 PheAsnAlaIleGluThrLeuProCysValLysLysLysAlaAspTrpAlaLeuHisTrp 210
QY 526 ATAGCAGATAGAAAATCTACTTTTGGGAAAGAGTGGTGGCTTGTGCTGTAGAAGGA 585
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 211 IleGlyAspLysGlnAlaThrPheGlyGluArgValValAlaPheAlaValGluGly 230
QY 586 GTTTCTCTCAGGATCTTTGCTGCTATATCTTGGCTAAAGAGAGAGTCTTATGCCA 645
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 231 IlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArgGlyLeuMetPro 250
QY 646 GGACTCATTCTTCCAAATGACATCATCAGCAGAGATTCAGAGCTTCTACTGTGCTTGTGCT

Db 251 GlyLeuThrPheSerAsnGluLeuLeuSerArgAspGluGlyLeuHisCysAspPheAla 270
QY TGCCTGATGTTCCATACTAGTAATAAGCTTCAGAGAAAGGTCAGGAGATCATTT 765
Db 271 CysLeuMetPheLeuHisLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 290
QY 766 GTTGATGCTGCAAAATTCAGCAGGAGTTCCTTAAACAGAGCCTCCAGGCTTCATT 825
Db 291 ThrAspAlaValGlnLeuGluGlnGluPheLeuThrGluAlaLeuProValAsnLeuLeu 310
QY 826 GGAATGAATTCGATTCGATGAACACAGTACATGATGATGTTAGCTGACAGATTCATTG 885
Db 311 GlyMetAsnCysThrLeuMetLysGlnTrpIleGluPheValAlaAspArgLeuLeuLeu 330
QY 886 GAACCTTGATTCCTCAAGGTTTCCTCAGCAGAAATCCTTTGATTTTATGGAACATTT 945
Db 331 GluLeuGlyPheLeuLysValPheLysAlaThrAsnProPheAspPheMetGluAsnLeu 350
QY 946 TCTTTAGAGGAAACAAATTCCTTTGAGAACAGGTTTCAGAGTATCAGCGTTTTCGA 1005
Db 351 SerLeuGluGlyLysThrAsnPhePheGluLysValGlyGluTyGlnLysMetGly 370
QY 1006 GTTATGGCAGAAACACAGATACCTCTTCACCTTGGATGGAGTTT 1053
Db 371 ValMetSerLysAlaLysAspAsnThrPheThrLeuAspAlaAspPhe 386

RESULT 12
RIR2_BRARE ID RIR2_BRARE STANDARD; PRT; 386 AA.
AC P79733;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Ribonucleoside-diphosphate reductase M2 chain (EC 1.17.4.1)
DE (Ribonucleotide reductase protein R2 class I).
GN Name=r2m2; Synonym=r2;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97137859; PubMed=8983196;
RA Mathews C.Z., Sjoberg B.-M., Karlsson M.;
RT "Cloning and sequencing of cDNAs encoding ribonucleotide reductase
from zebrafish Danio rerio."
RL Mol. Mar. Biol. Biotechnol. 5:284-287(1996).
CC -!- FUNCTION: Provides the precursors necessary for DNA synthesis.
CC -!- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate +
thioredoxin disulfide + H(2)O = ribonucleoside diphosphate +
thioredoxin.
CC -!- COFACTOR: Binds 2 iron ions per subunit (By similarity).
CC -!- PATHWAY: DNA replication pathway; first step.
CC -!- SUBUNIT: Heterodimer of a large (M1) and a small chain (M2).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the ribonucleoside diphosphate reductase
small chain family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
or send an email to license@isb-sib.ch).

DR EMBL; U57965; AB337103.1; -
DR HSPB; P11157; 1XSM.
DR ZFIN; ZDB-GENE-990415-25; rrm2.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonucl_redctse.

DR Pfam; PF00268; Ribonuc_red_sm; 1.
KW PROSITE; PS00368; RIBORED_SMALL; 1.
FT METAL 135 135 Iron 1 (By similarity).
FT METAL 166 166 Iron 1 and 2 (By similarity).
FT METAL 169 169 Iron 1 (By similarity).
FT METAL 229 229 Iron 2 (By similarity).
FT METAL 263 263 Iron 2 (By similarity).
FT METAL 266 266 Iron 2 (By similarity).
FT ACT_SITE 173 173 By similarity.
SQ SEQUENCE 386 AA; 44594 MW; C168846FB57F9F4E CRC64;

Alignment Scores:
Pred. No.: 2,33e-116 Length: 386
Score: 1425.50 Matches: 268
Percent Similarity: 89.41% Conservative: 36
Best Local Similarity: 78.82% Mismatches: 31
Query Match: 77.22% Indels: 5
DB: 1 Gaps: 1

US-10-698-228-2 (1-1053) x RIR2_BRARE (1-386)

QY 49 AGATCATCTTCAGACACCAACGAAAGTGAATAAAGTCA-----AATGAA 93
Db 47 ArgLysIlePheAspGluSerGluGlyGlnSerLysAlaLysLysGlyAlaValGluGlu 66
QY 94 GAGCCACTCTAAGAAAGAGTTCTCGCCGTTTGTCTATCTTCCAAATCCAGTACCCCTGAT 153
Db 67 GluProLeuLeuLysGluAsnProHisArgPheValIlePheProIleGlnTrpHisAsp 86
QY 154 ATTTGGAAATGTATAAACACAGGCACAGGCTTCTCTGGACAGCAAGAGGTCGACTTA 213
Db 87 IleTrpGlnMetTyLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAspLeu 106
QY 214 TCAAGGATCTCCCTCACTGGAACAAGCTTAAGCAGATGAGAGTACTTCACTCTCAC 273
Db 107 SerLysAspLeuGlnHisTrpAspSerLeuLysAspGluGluArgTyPheIleSerHis 126
QY 274 ATCTTAGCTCTTTTGCAGCCAGTGTGAATTTGTAATGAAATTTGGTGGAGCGCTTT 333
Db 127 ValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArgPhe 146
QY 334 AGTCAGGAGGTGCAGGCTCCAGAGGCTCGCTGTTCTATGGCTTTCAAATTTCTCATCAG 393
Db 147 ThrGlnGluValGlnValThrGluAlaArgCysPheTyGlyPheGlnIleAlaMetGlu 166
QY 394 ATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCAAGAA 453
Db 167 AsnIleHisSerGluMetTySerLeuLeuIleAspThrTyIleLysAspSerLysGlu 186
QY 454 AGGGAATTTTATTAAATGCAATTGAAACCATCGCCCTATGTTAAAGAAAAAGCAGATTGG 513
Db 187 ArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAspTrp 206
QY 514 GCCTTTCGATGATAGCAGATAGCAATAATCTACTTTTGGGAAAGAGTGGGCTTTGCT 573
Db 207 AlaLeuAsnTrpIleGlyAspLysAsnAlaArgTyGlyGluArgValAlaPheAla 226
QY 574 GCTGTAGAAAGAGTTCCTCTCAGGATCTTTTGTGCTATATTTCTGGCTTAAGAAAGAGA 633
Db 227 AlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArg 246
QY 634 GGTCTTATGCCAGGACTCACTTTTCCATGAATCACTCATCAGCAGATGAAGACTTCAC 693
Db 247 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHis 266
QY 694 TGTGACTTTTGTGCTGATGTTTCCAAATCTTAGTAAATTAAGCCTTCAGAAAGAGGCTC 753
Db 267 CysAspPheAlaCysLeuMetPheLysHisLeuIleAsnLysProSerGluGluThrVal 286
QY 754 AGGAGATCATTTGTGATGCTGTCAAAATTCAGCAGAGGTTTTTAAACAGAGCTTCCCA 813
Db 287 LysLysIleIleMetAsnAlaValGluIleGluGlnGluPheLeuThrAspAlaLeuPro 306

AC QSP876; 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 GN Names=MGC75900;
 OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8364;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gumaracne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S., Gerhard D.S.;
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
 DR ENBL, BC061353; AH61353.1; -
 DR GO; GO:0004748; F:ribonucleoside-diphosphate reductase activity; IEA.
 DR GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.
 DR InterPro; IPR009078; Ferritin/RR like.
 DR InterPro; IPR000358; Ribonucleotide reductase.
 DR Pfam; PF00268; Ribonucleotide reductase.
 DR PROSITE; PS00368; RIBORED_SMALL; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 378 AA; 43939 MW; 538D477FBB7CDB65 CRC64;
 Alignment Scores:
 Pred. No.: 4,47e-114 Length: 378
 Score: 1399.50 Matches: 262
 Percent Similarity: 89.77% Conservative: 45
 Best Local Similarity: 76.61% Mismatches: 28
 Query Match: 75.81% Indels: 7
 DB: 2 Gaps: 1

US-10-698-228-2 (1-1053) x Q6P876 (1-378)

QY 49 AGATCATCTTCAGACACCAAGTGAATAA-----AAGTCA 87
 Db 37 ArgThrSerArgGlyThrHisLysGlnTrpValCysGlnSerLeuLysAspProArgIle 56
 QY 88 AATGAAGACCTCTCTAAGAAAGATTCTCGCGGTTTGTCTATCTTTCAATCCAGTAC 147
 Db 57 GlnAspGluProLeuLeuArgAspAsnProGlyArgPheValIleLeuProIleGluTyr 76
 QY 148 CCTGTATTTGGAAATGTATTAACAGGACAGGCTTCTCTCGACACAGGAGGTC 207
 Db 77 HisAspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluVal 96
 QY 208 GACTTATCAAGGATCTCCCTCACTCGAACAAGCTTAAAGCAGATGAGAAGTACTTCATC 267

Db 97 AspLeuSerLysAspLeuProHisTrpGluAlaLeuLysProGluGluArgTyrPheIle 116
 QY 268 TCTCACATCTTAGCCTTTTTCGACCCAGTGTAAATGTAATGAAATTTGGTGGAG 327
 Db 117 SerTyrValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGlu 136
 QY 328 CGCTTTAGTCAGGAGTGCAGGTTCCAGAGCTCGCTCTTTCTATGGCTTTCAATTC 387
 Db 137 ArgPheSerGlnGluValGlnValThrGluAlaArgCysPheTyrGlyPheGlnIleAla 156
 QY 388 ATCGAGAATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACATTACATCAGATCCC 447
 Db 157 MetGluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspPro 176
 QY 448 AAGAAAAGGAATTTTATTTAATGCAATTCAAACCATGCCCTATGTTAAAGAAAAAGCA 507
 Db 177 LysGluArgGluPheLeuPheAsnAlaIleGluThrLeuProCysValLysLysAla 196
 QY 508 GATTGGCCCTTGCATGGATAGATAGATAAGAAATCTACTTTGGGGAAGAGTGTGGCC 567
 Db 197 GluTrpAlaLeuArgTrpIleSerAspArgGluAlaProPheCysGluArgIleValAla 216
 QY 568 TTTGCTCTGTAGAGAGTTTCTCTCAGATCTTTTGTCTGCTATATTTCTGGCTAAG 627
 Db 217 PheAlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLys 236
 QY 628 AAGAGAGGTCTTATGCCAGGACTCCTTTTCCATGAACTCACTCAGCAGAGATCAAGGA 687
 Db 237 LysArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGly 256
 QY 688 CTTCACTGTCACTTGTCTGCTGCTGCTCAATCTAGTAAATAGCCCTTCAGAAGAA 747
 Db 257 LeuHisThrAspPheAlaCysLeuLeuPheLysHisValValHisLysProSerGluGln 276
 QY 748 AGGTCAGGAGATCATTTGTTGATGCTGCTCAAAATTGACAGAGAGTTTTAAACAGAAGCC 807
 Db 277 ArgValArgGluIleIleAspAlaValLysIleGluGlnGluPheLeuThrGluAla 296
 QY 808 TTGCCAGTTGGCTCATTTGGAATGCAATGCAATTTGATGAACAGATACATTGAGTTTGA 867
 Db 297 LeuProValGluLeuIleGlyMetAsnCysLysLeuMetLysGlnTyrIleGluPheVal 316
 QY 868 GCTGACAGATTACTTGTGGAACCTTGGATTCTCAAGGTTTTCAGGCAGAGAAATCCTTT 927
 Db 317 AlaAspArgLeuLeuLeuGluGlyPheLysLysIlePheLysSerGluAsnProPhe 336
 QY 928 GATTTATGGAACAATTTCTTTAGAAAGGAAAAACAATTTCTTTGAGAAAACGAGTTTCA 987
 Db 337 AspPheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysLysValAla 356
 QY 988 GAGTATCAGCGTTTTCAGTATTGCGAGAAACACAGATAACAGTCTTCCCTTCATGATGCA 1047
 Db 357 GluTyrGlnLysMetGlyValMetSerLysAlaThrAspLysGlnPheThrLeuAsnAla 376
 QY 1048 GATTTT 1053
 Db 377 GluPhe 378

Search completed: October 30, 2005, 06:33:51
 Job time : 89 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 30, 2005, 06:01:36 ; Search time 15.5 Seconds

(without alignments)
10142.641 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1846

Sequence: 1 atggggaccggaagcc.....tcaccttgatgcagatttt 1053

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=framer_n2p.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US10698228/runat 26102005 100609 4010/app query.fasta_1.2446
-DB=Issued Patents AA -QFMT=fastan -SUPFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40:cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228 @CGN 1 1 32 @runat 26102005 100609 4010 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1821	98.6	351	4	US-10-019-733-1
2	1454	78.8	389	4	US-10-019-733-5
3	1454	78.8	389	4	US-09-949-016-6016
4	1454	78.8	453	4	US-09-949-016-7896
5	1131	61.3	411	4	US-09-248-796A-18017
6	1072.5	58.1	274	1	US-08-307-499-19
7	1072.5	58.1	274	3	US-09-299-268-19
8	974.5	52.8	322	1	US-08-136-743B-2
9	969.5	52.5	360	4	US-09-248-796A-18018
10	768	41.6	147	3	US-08-905-223-358
11	599	32.4	149	4	US-09-270-767-46236
12	495	26.8	323	4	US-09-902-540-10006

13	381.5	20.7	445	4	US-09-328-352-4714	Sequence 4714, Ap
14	377.5	20.4	476	4	US-09-252-991A-19358	Sequence 19358, A
15	335	18.1	365	4	US-09-198-452A-1059	Sequence 1059, Ap
16	335	18.1	380	4	US-09-438-185A-986	Sequence 986, Ap
c 17	299.5	16.3	121	1	US-08-307-499-7	Sequence 7, Appl
c 18	299.5	16.3	121	3	US-09-299-268-7	Sequence 7, Appl
19	298	16.1	314	3	US-09-080-044-3	Sequence 3, Appl
20	298	16.1	314	3	US-09-531-857A-3	Sequence 3, Appl
21	294	15.9	56	3	US-08-905-223-362	Sequence 362, App
22	220	11.9	382	4	US-09-543-681A-4239	Sequence 4239, Ap
23	203.5	11.0	412	4	US-09-489-039A-8180	Sequence 8180, Ap
24	201.5	10.9	401	4	US-09-540-236-3566	Sequence 3566, Ap
25	200	10.8	325	4	US-09-107-532A-4443	Sequence 4443, Ap
26	195.5	10.6	320	3	US-09-092-437-2	Sequence 2, Appl
27	195.5	10.6	320	4	US-09-583-110-3858	Sequence 3858, Ap
28	195.5	10.6	320	4	US-09-107-433-3136	Sequence 3136, Ap
29	195	10.6	339	4	US-09-543-681A-5069	Sequence 5069, Ap
30	169.5	9.2	323	4	US-09-042-775-2	Sequence 2, Appl
31	167	9.0	343	3	US-09-134-001C-5484	Sequence 5484, Ap
32	163	8.8	334	4	US-09-134-000C-5726	Sequence 5726, Ap
33	161	8.7	311	4	US-09-710-279-1260	Sequence 1260, Ap
34	157	8.5	324	2	US-08-813-940-2	Sequence 2, Appl
35	156	8.5	391	4	US-09-489-039A-10121	Sequence 10121, A
36	153	8.3	104	4	US-09-230-485-5	Sequence 5, Appl
37	138	7.5	321	2	US-08-813-940-6	Sequence 6, Appl
38	127	6.9	52	3	US-08-917-299-21	Sequence 21, Appl
39	127	6.9	52	3	US-09-422-662-21	Sequence 21, Appl
40	127	6.9	52	4	US-09-730-763-21	Sequence 21, Appl
41	127	6.9	52	4	US-09-429-370-21	Sequence 32, Appl
42	103	5.6	555	3	US-09-173-151A-32	Sequence 32, Appl
43	100.5	5.4	382	4	US-09-134-000C-5889	Sequence 5889, Ap
44	99	5.4	233	3	US-08-993-825-2	Sequence 2, Appl
45	98.5	5.3	896	4	US-09-270-767-46130	Sequence 46130, A

ALIGNMENTS

RESULT 1
US-10-019-733-1
; Sequence 1, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 1
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-1

Alignment Scores:
Pred. No.: 4,89e-211 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 4 Gaps: 0

US-10-698-228-2 (1-1053) x US-10-019-733-1 (1-351)

QY

1 ATGGGCGACCCGGAAGCGCGGCTGGATCAGATGAGATCATCTTCA 60

```
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAAGGAAAGTAAAGTCAATGAAGCCACTCTTAAGAAGAGTTCTCCG 120
Db 21 AspThrAsnGluSerGluIleYsSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGTTTGTGCATCTTCCCAATCAGTACCTCGATATTGGAATGTATATAACAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GTTCTCTTGTGACAGCAGAAAGAGTTCGACTTATCAAGGATCTCCCTCAGTGAACAAG 240
Db 61 AlaSerPheTrpTrpAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAGCAGATGAGAGTACTTCTCTCATCTCTACATCTTACCTTTTTCGCCAGTGAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTGTAATAAGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValProGluAla 120
QY 361 CGCTGTTCTATGCGTTTCAAAATCTCATPCGAGAATGTTCTACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGATCCCAAGAAAGGAATTTTATTAATGCAATGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAAGAAAAAGCAGATTGGCGCTTCGATGGATACAGATAGAAA 540
Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTGTCTGTGTAAGAAGAGTTTTCTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValIleGluValPhePheSerGly 200
QY 601 TCTTTTGTGCTATATTCTGGCTTAAGAAGAGAGGTCTTATGCCAGGACTCATTTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTCATCAGCAGAGATCAAGACTTCACTGTGACTTGTCTGCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAAGCCTTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db 241 TyrLeuValAsnLysProSerGluArgValArgGluIleValAspAlaValLys 260
QY 781 ATGAGCAGGAGTTTTAAACAGAGCCTTGGCAGTTGGCCTCATTTGGAATGCAATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACACTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTGCACTCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTTCAGCAGAAATCCTTTTGATTTTATGGAATAACATTTCTTTAGAGGAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluLys 320
QY 961 ACAAAATTTCTTGAAACACAGTTTTCAGAGTATCAGCGTTTTCAGATTTAGCAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTTTCACCTTGAATGCAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 2

US-10-019-733-5

```
; Sequence 5, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 5
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-5

Alignment Scores:
Pred. No.: 1,27e-166 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 4 Gaps: 0

US-10-698-228-2 (1-1053) x US-10-019-733-5 (1-389)
```

```
QY 91 GAAGAGCCACTCTCTAAGAAAGAGTTCTCGCGGTTTGTCTCTTCCATCCAGTACCT 150
Db 69 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 88
QY 151 GATATTGGAATAATGATATAACAGGCACAGCTTCTCTCGACAGCAGAGAGGTCCGAC 210
Db 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
QY 211 TTATCAAAAGGATCTCCCTCACTCGAACCAAGCTTAAAGCAGATGAGAGTACTTCTCT 270
Db 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTTAGCCTTTTTCGACGAGTGATGGAATTTGTAATGAAATTTTGGTGGAGCGC 330
Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY 331 TTTTAGTCAGAGGTGAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTTCTCATC 390
Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGAATGTTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCAG 450
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAAGGGAATTTTATTAATGCAATTTGAAACCAACGATCCCTATGTTAAGAAAAAGCAGAT 510
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
QY 511 TGGGCTTTCGAGTGGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCCTTT 570
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaAlaPhe 228
QY 571 GCTGCTGTAGAAGAGTTCCTCTCAGGATCTTTTCTGCTATATCTGCTGCTAAAGAG 630
Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGTCTTATGTCAGGACTCAGTCTTTTCCAATGAACATCAGCAGAGATGAAGGACTT 690
Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
```

QY 691 CACTGTGACCTTGGCTGCTGATGTTCCCAATCTAGTAAATAGCCTTCAGAGAAAGG 750
 Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
 QY 751 GTCAGGAGATCATTTGCTGCTGCTCAAAATTTGACAGGAGTTTAAACAGAGCCTTG 810
 Db 289 ValArgGluLeuLeuLeuLeuAlaValArgGluGluGlnGluPheLeuThrGluAlaLeu 308
 QY 811 CCAGTTGGCTCATTTGGAATGAATTCATTTGATGAAACAGTACATTTGATTTGAGCT 870
 Db 309 ProValLysLeuLeuLeuGlyMetLysCysThrLeuMetLysGlnTyrLeuGluPheValAla 328
 QY 871 GACAGATTACTTGTGGAATCTGGAATCTCAAGTTTTCAGCAGAGAAATCTTTTGAT 930
 Db 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
 QY 931 TTTATGGAACATTTCTTTAGAGAAACAAATCTTTTCAGAGAAACGAGTTTCAGAG 990
 Db 349 PheMetGluAsnLeuSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
 QY 991 TATCAGCGTTTGCAGTTATGCGAGAAACACAGATAACGCTCTTACCTTGGATGCAGAT 1050
 Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
 QY 1051 TTT 1053
 Db 389 Phe 389
 RESULT 3
 US-09-949-016-6016
 ; Sequence 6016, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949, 016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 6016
 ; LENGTH: 389
 ; TYPE: PRT
 ; ORGANISM: Human
 ; US-09-949-016-6016
 Alignment Scores:
 Pred. No.: 1,27e-166 Length: 389
 Score: 1454.00 Matches: 269
 Percent Similarity: 94.08% Conservative: 33
 Best Local Similarity: 83.80% Mismatches: 19
 Query Match: 78.76% Indels: 0
 DB: 4 Gaps: 0
 US-10-698-228-2 (1-1053) x US-09-949-016-6016 (1-389)
 QY 91 GAAGAGCCACTCTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTCCAAATCCAGTACCCT 150
 Db 69 AspGluProLeuLeuArgGluAsnProArgPheValLeuPheProLeuGluTyrHis 88
 QY 151 GATATTGAAATGATATAACAGGACAGGCTTCTTTCGACAGCAGAGAGGTCGAC 210
 Db 89 AspLeuTrpGlnMetTyrLysLysAlaGluAlaSerPheThrAlaGluGluValAsp 108
 QY 211 TTATCAAGAGTCTCCCTCAGTGAACAGCTTAAAGCAGATGAGAGTACTTCTCTCT 270

Db 109 LeuSerLysAspLeuGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
 QY 271 CACATCTTAGCTTTTTCAGCCAGTATGGAATTTGAAATGAAATTTTGTGTGAGCCG 330
 Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyLeuValAsnGluAsnLeuValGluArg 148
 QY 331 TTTAGTCAGGAGGTGTCAGGCTCCAGAGCTCGCTGTTTCTATGCTTTCAAAATTCATC 390
 Db 149 PheSerGlnGluValGlnLeuThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
 QY 391 GAGATGTTTCTACTCAGAGATGTACAGTTTCTGCTGATAGACACTTACATCAGAGATCCCAAG 450
 Db 169 GluAsnLeuHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
 QY 451 AAAAGGGAATTTTATTAATGCAATTAACCAATGCAATGCTATGTTTAAAGAAAACAGAT 510
 Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
 QY 511 TGGGCTTTCAGATGATGATAGAGATAGAAATCTACTTTTGGGGAAGAGTGTGGCCTTT 570
 Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 228
 QY 571 GCTGCTTAGAGAGATTTTCTTCTCAGATCTTTTCTGCTGCTATATTCTGCTAAAGAG 630
 Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
 QY 631 AGAGTCTTATGCCAGGACTCAGTTTTCCTCAATCACTCACTCAGCAGAGATGAAGACTT 690
 Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
 QY 691 CACTGTGACTTTCCTTGCCTGATGTTCCAAATCTAGTAAATAGCCTTCAGAGAAAGG 750
 Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
 QY 751 GTCAGGAGATCATTTGTTGATGCTGTCAAAATTCAGCAGAGAGTTTTAAACAGAGCCTTG 810
 Db 289 ValArgGluLeuIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
 QY 811 CAGTTGGCTCATTTGGAATGAATTCATTTGATGATGAAACAGTACATTTGATTTAGCT 870
 Db 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
 QY 871 GACAGATTACTTGTGGAATTCCTCAAGGTTTTCAGGCGAGAAATCCTTTTTCAT 930
 Db 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
 QY 931 TTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTTCAGAG 990
 Db 349 PheMetGluAsnLeuSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
 QY 991 TATCAGCGTTTTCAGTTATGCGAGAAACACAGATTAACGCTTTCACCTTGGATGCAGAT 1050
 Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
 QY 1051 TTT 1053
 Db 389 Phe 389
 RESULT 4
 US-09-949-016-7896
 ; Sequence 7896, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949, 016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR FILING DATE: 2000-10-03


```
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7896
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7896

Alignment Scores:
Pred. No.: 1.38e-166 Length: 453
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 4 Gaps: 0

US-10-698-228-2 (1-1053) x US-09-949-016-7896 (1-453)
QY 91 GAAGAGCCACTCTAAGAAGAGTTCTCGCGGTTTGTCTATCTTCCAAATCCAGTACCCT 150
DB 133 AspGluProLeuMetGluAsnProArgPheValIlePheProIleGluThrHis 152
QY 151 GATATTGAAATGTATAACAGGCACAGGCTTCTCTGGACAGCAGAGGTCCGAC 210
DB 153 AspileTIpGluMetTyrlsYsAlaGluAlaSerPheThrAlaGluGluValAsp 172
QY 211 TTATCAAGGATCTCCCTCACTGGAACAGCTTAAGCAGATGAGAACTTCATCTCT 270
DB 173 LeuSerLysAspileGlnHisTrpGluSerLeuLysProGluGluArgTyrlsPheIleSer 192
QY 271 CATCTTGTAGCTTTTTCACGCGCTGATGAAATGTAATGAAATTTGGTGGAGCG 330
DB 193 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 212
QY 331 TTATGTCAGGAGTGCAGGCTTCAGAGCTCGCTGTTTCTATGGCTTTCAAAATCTCATC 390
DB 213 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrlsGlyPheGlnIleAlaMet 232
QY 391 GAGATGTTCACTCAGAGATGACATGTTGCTGATGACACTTACATCAGAGATCCCAAG 450
DB 233 GluAsnIleHisSerGluMetTyrlsSerLeuLeuIleAspThrTyrlsYsAspProLys 252
QY 451 AAAAGGGAATTTTATTTAATCCAAATGAAACCATCCCTATGTTTAAAGAAACAGAT 510
DB 253 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 272
QY 511 TGGGCTTTCGATGATAGATAGAAATCTACTTTTGGGAAAGAGTGGTGGCTTT 570
DB 273 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrlsGlyGluArgValAlaAlaPhe 292
QY 571 GCTGCTGTAGAAGGATTTCTCTCAGGATCTTTTGTGCTATATTTGGCTAAAGAG 630
DB 293 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 312
QY 631 AGAGTCTTATGCGAGGACTCATTTTTCATGAACTCATGCGAGAGATGAGGACTT 690
DB 313 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 332
QY 691 CACTGTGACTTTGCTGCTGATGTTTCCAAATCTAGTAAATAAGCCTTTCAGAAAGAG 750
DB 333 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 352
QY 751 GTCAGGAGATCATTTGATGCTGTCAAATTTGACGAGGATTTTAAACAGAGCTTTG 810
DB 353 ValArgGluIleIleAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 372
QY 811 CAGTTGGCTCATTTGGAATGATTTGATTTGATTAATGAAACAGTACATTTGAGTTT 870
DB 373 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrlsIleGluPheValAla 392
QY 871 GACAGATTACTTGTGGAACCTTGGATTCTCAAGGTTTTCAGGCGAGAAATCCTTTTGAT 930
```

```
DB 393 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 412
QY 931 TTTATGAAAACATTTCTTTAGAGGAAAACAAATTTCTTTGAGAAACAGTTTCAGAG 990
DB 413 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 432
QY 991 TATCAGCGTTTTCAGTTATGCGAGAAACCCACAGATAACGCTTTCACCTTGGATGCGAGAT 1050
DB 433 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 452
QY 1051 TTT 1053
DB 453 Phe 453

RESULT 5
US-09-248-796A-18017
; Sequence 18017, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18017
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18017

Alignment Scores:
Pred. No.: 1.54e-127 Length: 411
Score: 1131.00 Matches: 210
Percent Similarity: 78.21% Conservative: 52
Best Local Similarity: 62.69% Mismatches: 67
Query Match: 61.27% Indels: 6
DB: 4 Gaps: 3

US-10-698-228-2 (1-1053) x US-09-248-796A-18017 (1-411)
QY 67 AACGAAAGTGAATAAGTCAAAATGAA-----GAGCCACTCTCTAAGAAAGAGTTCTCGC 120
DB 77 AsnArgHisLysLeuLysGlnGluSerAsnGluProLeuLeuThrProAspLysThr 96
QY 121 CGTTTGTTCATCTTTCCAAATCCAGTACCTGATATTTGGAAAATGTATATAACAGGCACAG 180
DB 97 ArgHisThrIleTyrlsProIleLysTyrlsProGluLeuTrpGlnPheTyrlsLysSerLeu 116
QY 181 GCTTCCTTTTGGACAGCAGAGGTCGACTTATCAAGGATCTCCCTCACTGG---AAC 237
DB 117 AlaSerPheTrpThrAlaGluLeuAspLeuSerLysAspLeuAspAspTrpAsnAsn 136
QY 238 AAGCTTAAAGCAGATGAGAACTTCTCATCTCTCATCTTTAGCTTTTTCAGCCAGT 297
DB 137 LysMetAsnGluAsnGluArgPhePheIleSerArgValLeuAlaPhePheAlaAlaSer 156
QY 298 GATGGAATGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCCAGAG 357
DB 157 AspGlyIleValAsnGluAsnLeuValGluAsnPheCysAlaGluValGlnIleProGlu 176
QY 358 GCTCGCTGTTTCTATGGCTTTCAAAATTCATCAGAAATGTTCACTCAGAGATGACAGT 417
DB 177 AlaLysLeuValTyrlsPheGlnIleMetMetGluAsnIleHisSerGluThrTyrlsSer 196
QY 418 TTGCTCATAGACACTTATCATCATGAGATCCCAAGAAAGGAAATTTTATTATTAATGCAATT 477
```

```
Db 197 LeuLeuLeuGluThrTyrPheLysAspProGluGluAlaAspPheLeuPheAsnAlaIle 216
QY 478 GAACCATGCCCTATGTTAAGAAAAACAGAGATTGGCGCTTGGATGATAGCAGATAGA 537
Db 217 AspAsnIleProPheIleArgLysLysAlaAspTrpAlaIleArgTrpIleGlnSerGlu 236
QY 538 AATCTACTTTTGGGGAAGAGTGGCGCTTGGCTGCTGCTAGAGAGGATTTCTCTCA 597
Db 237 AspAlaLeuTyrAlaGluArgLeuValAlaPheAlaValGluGlyIlePhePheSer 256
QY 598 GGATCTTTTGGCTGCTATCTGCTTAAAGAAAGAGAGGCTTATGCCAGGACTCATT 657
Db 257 GlySerPheAlaIlePheTrpLeuLysArgGlyLeuMetProGlyLeuThrPhe 276
QY 658 TCAATGAACTCATCAGCAGATGAAGGACTTCACTGTGCTGCTGCTGCTGCTGCTGCT 717
Db 277 SerAsnGluLeuLeuCysArgAspGluGlyIleHisThrAspTyrAlaCysLeuPhe 296
QY 718 CAATACTAGTAATAAGCCTTCAGAAAGAGGCTCAGGAGATCAATTGTTGATGCTGTC 777
Db 297 SerTyrLeuLysAsnLysProSerSerGluIleLeuLysIleIleThrGluAlaVal 316
QY 778 AATATTGACGAGAGTTTAAACAGAGCCTTCCAGTGGCTCATTTGGAATGCAATGC 837
Db 317 AspIleGluLysGluTyrPheSerAspAlaLeuProValSerLeuLeuGlyMetAsnCys 336
QY 838 ATTTGATGAAACAGTACATGAGTTGAGCTGACAGATTACTTGTGGAAGTGGATTC 897
Db 337 AspLeuMetCysGlnTyrValGluPheValAlaAspArgLeuLeuValAlaPheGlyAsn 356
QY 898 TCAAGCTTTTTCAGGCAGAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGGA 957
Db 357 LysLysTyrTyrAsnValThrAsnProPheAspPheMetGluAsnIleSerLeuAlaGly 376
QY 958 AAAACAAATTTCTTTCAGAAACAGATTTCAGATATCAGCGTTTTCAGTATGCGAGAA 1017
Db 377 LysThrAsnPhePheGluLysArgValSerAspTyrGlnLysAlaGlyValMetGluLys 396
QY 1018 ACC-----ACAGTACGCTCTTCACCTTGATGATGATGATGATGATGATGATGAT 1053
Db 397 ValGluAsnLysAsnGluLysThrGlyLeuPheAspGlnAspPhe 411
```

RESULT 6

```
US-08-307-499-19
; Sequence 19, Application US/08307499
; Patent No. 5651972
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSES: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,499
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF35.1.FWCCI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 274 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-307-499-19
```

```
Alignment Scores:
Pred. No.: 1,48e-120 Length: 274
Score: 1072.50 MatChes: 204
Percent Similarity: 85.56% Conservative: 27
Best Local Similarity: 75.56% Mismatches: 38
Query Match: 58.10% Indels: 1
DB: 1 Gaps: 1
US-10-698-228-2 (1-1053) x US-08-307-499-19 (1-274)
```

```
QY 94 GAGCCTCTCTAAGAAAGAGTCTCGCGGTTTGTCTATCTTCCATCCAGTACCCCTGAT 153
Db 2 GluProIleLeuGlnGluSerAspSerArgPheValIlePheProIleLysTyrHisAsp 21
QY 154 ATTTGGAATATGATAACAGCAGCAGCTTCTCTGACAGCAGAGAGAGTCCACTTA 213
Db 22 IleTrpLysMetTyrLysGlnSerValAlaSerPheTrpThrValGluGluValAspLeu 41
QY 214 TCAAGGATCTCCCTCTCTGAAAGCTTAAAGCAGATGAGAACTTCTCTCTCTCAC 273
Db 42 SerLysAspLeuAspAspTrpAspLysLeuThrLysAspGluLysTyrPheIleLysHis 61
QY 274 ATCTTAGCCTTTTTCAGCCAGTGAATGTAATGAAATGAAATTTGTTGAGCGCTTT 333
Db 62 IleLeuAlaPhePheAlaSerSerAspGlyIleValAsnGluAsnLeuAlaGluArgPhe 81
QY 334 AGTCAGGAGTGCAGGTTCCAGAGCTCGCTTCTCTATGGCTTTCAAAATTTCTCATCGAG 393
Db 82 TyrValAspValGlnCysSerGluAlaArgCysPheTyrGlyPheGlnIleAlaMetGlu 101
QY 394 AATGTTTCACTCAGAGATGTACAGTTTGTGTATAGACACTTACATCAGAGATCCCAAGAA 453
Db 102 AsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrValArgAspAsnIleGlu 121
QY 454 AGGGAATTTTATTAATGCAATGAAACCATGCCCTATGTTAAGAAAAACAGCAGATTGG 513
Db 122 LysMetHisLeuPheAsnAlaIleGluThrMetGluCysValLysLysLysAlaAspTrp 141
QY 514 GCCTTGGATGGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCTTTTGGT 573
Db 142 AlaArgLysTrpIleSerSerAsnLys---ValTyrGlyGluArgValValAlaPheAla 160
QY 574 GCTGTAGAAGGAGTTTCTCTCTCAGGATCTTTTCTGCTATATCTGCTGCTGCTGCTGCT 633
Db 161 AlaValGluGlyIlePhePheSerGlySerPheAlaIleIlePheTrpIleLysArg 180
QY 634 GGTCTTATGCCAGGACTCATTCTTTTCCAATGAATCATCAGCAGAGATGAAGGACTTCA 693
Db 181 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHis 200
QY 694 TGTGACTTTGCTTGCCTGATGTTCCAACTTAGTAATAAGCCTTTCAGAGAAAGGGTTC 753
```

||||| 201 CysAspPheAlaCysLeuMetPheLysHisLeuLeuHisProSerLysGluValIle 220 :
QY 754 AGGGAGATCATTTGATGCTGCTCAAAATTAGCAGGAGTTTAAACAGAGCCTTGCCA 813 :
Db 221 ThrSerIleIleIleAspAlaValAsnIleGluLysGluPheLeuThrValAlaIlePro 240 :
QY 814 GTTGGCCTCATTTGGAATGAAATTTGATGAAACAGTACATTTGAGTTGTAGCTGAC 873 :
Db 241 ValAspLeuIleGlyMetAsnCysCysLeuMetSerGlnIryIleGluPheValAlaAsp 260 :
QY 874 AGATTACTTTGTGGAATTTGGATTCTCAAG 903 :
Db 261 ArgLeuLeuThrGluLeuGlyCysGluLys 270 :
RESULT 7

US-09-299-268-19
; Sequence 19, Application US/09299268
; Patent No. 6217882
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,268
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/901,127
; FILING DATE:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 274 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-299-268-19

Alignment Scores: 1.48e-120 Length: 274
Pred. No.: 1072.50 Matches: 204
Score:

Percent Similarity: 85.56% Conservative: 27
Best Local Similarity: 75.56% Mismatches: 38
Query Match: 58.10% Indels: 1
DB: 3 Gaps: 1
US-10-698-228-2 (1-1053) x US-09-299-268-19 (1-274)
QY 94 GAGCCACTCTTAAGAAAGAGTTCTCGCCGGTTTGTCTATCTTTTCCAAATCCAGTACCCCTGAT 153
Db 2 GluProIleLeuGlnGluSerAspSerArgPheValIlePheProIleLysTyrHisasp 21
QY 154 ATTTGGAAAATGTATAACAGCAGGCTTCCTTCTGGACAGCAGAGAGGTTCGACTTA 213
Db 22 IleTrpLysMetTyrLysGlnSerValAlaSerPheTrpThrValGluGluValAspLeu 41
QY 214 TCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTCTCTCAC 273
Db 42 SerLysAspLeuAspAspTrpAspLysLeuThrLysAspGluLysTyrPheIleLysHis 61
QY 274 ATCTTAGCCCTTTTTCAGCCAGTGTGAATTTGTAATGAAATTTGGTGGAGCGCTTT 333
Db 62 IleuAlaPhePheAlaSerSerAspGlyIleValAsnGluAsnLeuAlaGluArgPhe 81
QY 334 AGTCAGGAGGTGCAGGTTCAGAGGCTCGCTGTTCCTATGGCTTTCAAATTTCTCATCAG 393
Db 82 TyrValAspValGlnCysSerGluAlaArgCysPheTyrGlyPheGlnIleAlaMetGlu 101
QY 394 AATGTTCACTCAGAGATGATACAGTTTGTCTGATAGACACTTACATCAGAGATCCCAAGAAA 453
Db 102 AsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrValArgAspAsnIleGlu 121
QY 454 AGGGAATTTTATTGAATGCAATTTGAAACCATGCCCTATGTTAAGAAAAAGCAGATTGG 513
Db 122 LysMetHisLeuPheAsnAlaIleGluThrMetGluCysValLysLysLysAlaAspTrp 141
QY 514 GCCTTGGATGATGACAGATAGAAAATCTACTTTTGGGAAAAGAGTGGTGGCCTTTCCT 573
Db 142 AlaArgLysTrpIleSerSerAsnLys--ValTyrGlyGluArgValValAlaPheAla 160
QY 574 GCTGTAGAGGAGTTTCTTCTCAGGATCTTTGCTGCTCTATATTTCTGGCTAAAGAGAGA 633
Db 161 AlavaIgluglyIlePhePheSerGlySerPheAlaAlaIlePheTrpIleLysLysArg 180
QY 634 GGTCTTATGCCAGGACTCACTTTTCCAATGAATCATCAGCAGATGAAGGACTTCAC 693
Db 181 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHis 200
QY 694 TGTGACTTTTCTGCTGCTGATGTTCCAAATCTAGTAAATTAAGCCTTCAGAAAGAGGCTC 753
Db 201 CysAspPheAlaCysLeuMetPheLysHisLeuHisProProSerLysGluValIle 220
QY 754 AGGGAGATCATTTGATGCTGCTCAAAATTAGCAGAGAGTTTAAACAGAGCCTTGCCA 813
Db 221 ThrSerIleIleIleAspAlaValAsnIleGluLysGluPheLeuThrValAlaIlePro 240
QY 814 GTTGGCCTCATTTGGAATTTGCAATTTTTCAGAACAGTACATTTGAGTTGTAGCTGAC 873
Db 241 ValAspLeuIleGlyMetAsnCysCysLeuMetSerGlnIryIleGluPheValAlaAsp 260
QY 874 AGATTACTTGTGGAATTTGGATTCTCAAG 903
Db 261 ArgLeuLeuThrGluLeuGlyCysGluLys 270

RESULT 8
US-08-136-743B-2
; Sequence 2, Application US/08136743B
; Patent No. 5459063
; GENERAL INFORMATION:
; APPLICANT: Barry S. Cooperman, Harvey Rubin,
; APPLICANT: Jerome Salem, and Alison L. Fisher
; TITLE OF INVENTION: "plasmodium falciparum Ribonu-
; TITLE OF INVENTION: cleotide Reductase, DNA Sequences Therefor and Peptide Inhibit
; TITLE OF INVENTION: Thereof"


```
QY 127 GTCATCTTTCCAAATCCAGTACCCTCATATTTGGAAAAATGTATAAACAGGACAGGCTTCC 186
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
112 ValMetPheProIleArgTyrHisGluIleTyrAsnPheTyrLysLysAlaGluAlaSer 131
QY 187 TTCTCGACAGCAGAGAGAGTGCAGTATCAAAAGGATCTCCCTCAGTGG---AACAGCTT 243
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
132 PheTyrThrAlaGluGluIleAspLeuSerLysAspLeuAspAspTrpAsnAsnLysLeu 151
QY 244 AAAGCAGATGAGAACTACTCTCTCATCTCTACATCTAGCCTTTTTCGACGCGATGATGA 303
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
152 AsnGluAsnGluArgTyrPheIleSerArgValLeuAlaPheAlaAlaSerAspGly 171
QY 304 ATTGTAATGAAATTTTGTGTGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCTCGC 363
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
172 IleValGlyGluAsnLeuIleGluAsnPheSerThrGluValGlnLeuProGluAlaLys 191
QY 364 TGTTCATATGCTTCAAAATTCATCGAGATGTCAGATGTCAGATGTCAGATGTCAGATGTC 423
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
192 SerPheTyrGlyPheGlnIleMetMetGluAsnIleHisSerGluThrTyrSerLeuLeu 211
QY 424 ATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATCAATTTGAACCC 483
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
212 IleGluThrTyrIleLysAspProGlnGluAlaAspTyrLeuPheAsnAlaIleAlaAsn 231
QY 484 ATGGCTATGTTTAAAGAAAAGCAGATTCGGCTTCGATGATGATGATGATGATGATGATG 543
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
232 IleProCysIleGlnLysAlaAspTrpAlaIleLysTrpIleGlnAspAspGluAla 251
QY 544 ACTTTTGGGGAAGAGTGTGTCCTTGTCTGTAGAGGAGTTCCTTCTCAGGATCT 603
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
252 LeuTyrGlyGluArgLeuValAlaPheAlaAlaValGluGlyIlePhePheSerGlySer 271
QY 604 TTGTCTGCTATATTCGGCTTAAAGAGAGAGGTCCTTATGCCAGGACTCTCTTTTCCAA 663
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
272 PheAlaSerIlePheTyrLysLysArgGlyLeuMetProGlyLeuThrPheSerAsn 291
QY 664 GAATCATCAGCAGAGATGAAGGACTTCTACTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 723
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
292 GluLeuIleCysArgAspGluGlyLeuHisThrAspPheAlaCysLeuLeuPheAlaHis 311
QY 724 TTAGTAATAAGCTTCAGAGAAAGGTCAGGAGATCATTTGTCATGCTGTCAAAAT 783
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
312 LeuGlnAsnArgProSerProGluIleValGluArgIleIleThrGluAlaValAspVal 331
QY 784 GAGCAGGAGTTCCTTAAAGAGCCTTGGCCAGTTCAGTTCGCTCATTTGGAATGATGCT 843
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
332 GluLysGluTyrPheThrAspValLeuProValSerLeuLeuGlyLeuAsnCysAsnLeu 351
QY 844 ATGAAACAGTACATTCAGTTTCTA 867
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
352 MetCysProTyrIleGluPheLeu 359

RESULT 10
US-08-905-223-358
; Sequence 358, Application US/08905223
; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 358:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 147 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Brain
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: -101...-1
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 4.4
; OTHER INFORMATION: seq ISHILAFFAASDG/IV
; US-08-905-223-358

Alignment Scores:
Pred. No.: 7,19e-84 Length: 147
Score: 768.00 Matches: 147
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 41.60% Indels: 0
DB: 3 Gaps: 0

US-10-698-228-2 (1-1053) x US-08-905-223-358 (1-147)
QY 1 ATGGCGACCCGAAAGCGCGGAGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAAGTGAATAAAGTCAATCAAGACGACTCTCTAAGAAAGAGTCTCGC 120
DB 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSer 40
QY 121 CGGTTTGTCTATCTTCCAAATCCAGTACCCCTGATATTTGGAAAATGTATAAACAGCAG 180
DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAla 60
QY 181 GCTTCCTTCTGGACAGCAGAGGTCGACTTATCAAAAGGATCTCCCTCATCTGGAACAAG 240
DB 61 AlaSerPheTyrThrAlaGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
QY 241 CTTAAAGCAGATCAGAGTACTTCTCATCTCTCATCTTACCTTTTTCGACGAGCTCAT 300
DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAlaSerAsp 100
QY 301 GGAATTTGTAATGAAATTTTGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTCAAAATTTCTCATCAGAAATCTTCTACTCAGAGATGTACAGTTTG 420
DB 121 ArgCysPheTyrGlyPheGlnIleLeuIleLeuValAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATGACACTTACATCAGA 441
DB 141 LeuIleAspThrTyrIleArg 147

RESULT 11
```

US-09-270-767-46236
; Sequence 46236, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46236
; LENGTH: 149
; TYPE: PRT
; ORGANISM: *Drosophila melanogaster*
US-09-270-767-46236

Alignment Scores:
Pred. No.: 2e-63 Length: 149
Score: 599.00 Matches: 114
Percent Similarity: 85.91% Conservative: 14
Best Local Similarity: 76.51% Mismatches: 21
Query Match: 32.45% Indels: 0
DB: 4 Gaps: 0

US-10-698-228-2 (1-1053) x US-09-270-767-46236 (1-149)
QY 586 GTTTCCTCTCAGGATCTTTCTGCTATATCTGTGCTAAAGAGAGAGGTCTTATGCCA 645
Db 1 IIEPHEPHESErGlySerPheAlaSerIIEPHErPLeuLysLysArgGlyLeuMetPro 20
QY 646 GGACTCAGCTTTTCAATGAACTCATCAGCAGAGATGAAGGCTTCACCTGTGCTTTGCT 705
Db 21 GlyLeuThrPheSerAsnGluLeuLeuSerArgGlyGlyLeuHisCysAspPheAla 40
QY 706 TGCCTGATGTTCCAAATACCTAGTAATAAGCCTTCAGAAAGAGGCTCAGGAGATCAT 765
Db 41 ValLeuMetPheGlnHisLeuValGlnArgProLysArgGluArgIleLeuGluLeu 60
QY 766 GTTGATGCTGCAAAATGAGCAGAGGATTTTAAACGAAGCCCTGCGAGTGGCTCAT 825
Db 61 ArgAspAlaValAlaIleGluGlnPheLeuThrAspAlaLeuProValAsnLeu 80
QY 826 GGAATGAATTCATTTTCATGAAACAGTACATTTGATGTTGATGCTCAGATCTACTGTG 885
Db 81 GlyMetAsnCysAspLeuMetSerGlnIleGluPheValAlaAspArgLeuLeuVal 100
QY 886 GAACCTTGATTTCTCAAGGCTTTTCAGGCAGAGAAATCCCTTTTGAATTTATGAAACAT 945
Db 101 GluLeuGlyValGlyValIleThrAsnThrLysAsnProPheAsnPheMetGluMetIle 120
QY 946 TCTTTAGAGGAAACAAATTTCTTTGAGAACGAGTTCAGATATCAGCGTTTGTCA 1005
Db 121 SerLeuAspGlyLysThrAsnPheGluLysValGlyGluThrGlnArgMetGly 140
QY 1006 GTTATGGCAGAAACACACATACCTC 1032
Db 141 ValValSerAsnProLeuAspAsnVal 149

RESULT 12
US-09-540-10006
; Sequence 10006, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: *Myxococcus xanthus* Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883

; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10006
; LENGTH: 323
; TYPE: PRT
; ORGANISM: *Myxococcus xanthus*
US-09-540-10006

Alignment Scores:
Pred. No.: 1.15e-50 Length: 323
Score: 495.00 Matches: 113
Percent Similarity: 54.49% Conservative: 63
Best Local Similarity: 34.98% Mismatches: 125
Query Match: 26.81% Indels: 22
DB: 4 Gaps: 6

US-10-698-228-2 (1-1053) x US-09-540-10006 (1-323)

QY 145 TACCCTGATATTTGGAAAAATGTATAAACAGCAGCAGCGCTCTCTCTGGACAGAGAAG 204
Db 3 TyrProGlnPhePheGluMetTyrArgAsnAlaIleLysAsnThrTyrThrValGlu 22
QY 205 GTCGACTTATCAAGAGATCTCCCTCCTG--AACAGCTTAAAGCAGATGAGAGTAC 261
Db 23 ValAspPheSerThrAspLeuValAspLeuArgSerLysMetThrAspAlaGluArgHis 42
QY 262 TTCACTCTCATCTTAGCCTTTTTCACGCCAGTGAATGTAATGTAATGTAATGTAATG 321
Db 43 LeuIleHisArgLeuValAlaPhePheAlaThrGlyAspSerIleValGlyAsnAsnLeu 62
QY 322 TCGAGCGCTTGTAGTCAGAGGTCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTCAA 381
Db 63 ValLeuAsnLeuTyrLysHisLeuAsnAlaProGluAlaArgMetTyrLeuSerArgGln 82
QY 382 ATTCTCATCGAGAATGTTCTCATCAGAGATGTACAGTTTGTCTGATAGACACATTCATCAGA 441
Db 83 TyrGluGluAlaLeuHisValGlnPheTyrLeuThrLeuLeuAspThrTyrValPro 102
QY 442 GATCCCMAGAAAAGGGAATTTTATTAAATGCAATTTGAACCATGCCCTATGTTAAGAAA 501
Db 103 AspProAlaGluArgAlaMetAlaPheAlaAlaValAspAsnIleProSerIleGlnArg 122
QY 502 AAAGCAGATTGGGCTTGCATGCGATA-----GCAGATAGAAAATCTACT 546
Db 123 LysAlaGlnPheCysMetLysTrpMetAspSerIleGlnGlyLeuAspThrLeuGlnThr 142
QY 547 TTTGGGGAAGA-----GTGCTGCTTTTGCT--GCTGTAGAAGA 585
Db 143 LysAlaHisArgArgGlnPheLeuLeuAsnLeuLeuLeuLeuLeuLeuLeuLeuLeu 162
QY 586 GTTTTCTTCTCAGATCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCA 645
Db 163 LeuPhePhePheAlaAlaPheAlaTyrValTyrPheLeuArgSerLysGlyLeuLeuAsn 182
QY 646 GGACTCAGCTTTTCCCAATGAACTCATCAGCAGAGATGAAGGACTTCAGCTGTGCTTCT 705
Db 183 GlyLeuAlaAlaGlyThrAsnTrpValPheArgAspGluSerAlaHisMetAlaPheAla 202
QY 706 TGCCTGATGTTCCAATAC-----TTAGTAAATAAGCTTTCAGAA 744
Db 203 PheGluSerIleGlnValAlaAlaArgLysGluGluProAspLeuPheAspAlaGlnMetGlu 222
QY 745 GAAAGGCTCAGGAGATCATTTGTGATGCTCTCAAAATTTGACGAGGAGTCTTAAACAGAA 804
Db 223 ArgAspValValAlaMetLeuArgGluAlaValGluCysGluThrGlnPheAlaGlnAsp 242
QY 805 GCCTTCCAGTTCGCCCTCATTTGAATGAATTTGATGAAACAGTACATGAGTTT 864
Db 243 LeuLeuSerGlyGlyValMetGlyValMetGlyLeuSerValGlnAspMetArgGlyTyrLeuGluTyr 262
QY 865 GTAGCTGACAGATTACTTGTGGAACTTGGATTCTCAAAAGGTTTTCAGGCGAGAAATCCT 924
Db 263 ValAlaAspGlnArgLeuGlnMetLeuGlyMetThrProValPheGlyThrLysAsnPro 282

QY 925 TTTGATTTTATGGAAACATCTTTCTTAGAAGGAAACAAATTTCTTTTGAGAAACGAGTT 984
Db 283 LeuHisPheMetAspLeuGlnAspValGlnGluLeuThrAsnPheGluArgVal 302
QY 985 TCAGAGTATCAGCGTTTGGCAGTTATGGCAGAAACCCACAGATAAGCTTTCACCTTGGAT 1044
Db 303 SerSerTyrGlnValAlaValGlyValGlyAlaAlaThrAspValVal-----LeuAsp 320
QY 1045 GCAGATTTT 1053
Db 321 AlaAlaPhe 323

RESULT 13
US-09-328-352-4714
; Sequence 4714, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4714
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4714

Alignment Scores:
Pred. No.: 7,35e-37 Length: 445
Score: 381.50 Matches: 97
Percent Similarity: 49.57% Conservative: 76
Best Local Similarity: 27.79% Mismatches: 143
Query Match: 20.67% Indels: 33
DB: 4 Gaps: 10

US-10-698-228-2 (1-1053) x US-09-328-352-4714 (1-445)

QY 25 GCGGCGCGGTGGATCAGGATCAGAGATCATCTTCAGACACCAACGAAAGTGAATAAAG 84
Db 98 AlaProGlyLeuGluGluLeuMetGlyAla-----GlnArgValGln 112
QY 85 TCAATGAAGACCCATCTTAAGAAGAGTTCTCCCGGTTTGTCTATCTTTTCCAATCCAG 144
Db 113 ValAspAspLysAlaMetileAsnCysArgAlaAspLeuAsnGlnLeuValProPheLys 132
QY 145 TACCTGTATTTGGAAATGTATTAACAGGCACAGGCTTCTCTGGACAGCAGAGAG 204
Db 133 TyrGluTrpAlaTrpGlnLysTyrLeuAspGlyCysAlaAsnHisTrpMetProGlnGlu 152
QY 205 GTCGACTTATCAAGAGTCTCCCTCACTCG-----AACAGCTTAAAGCAGATGAG 255
Db 153 ValAsnMetAsnHisAspIleAlaLeuTrpLysSerGluAsnGlyLeuThrGluAspGlu 172
QY 256 AAGTACTTCTCTCACATCTTAGCCTTTTTCAGCCAGGATGGAATGTGAATGAA 315
Db 173 ArgThrIleValMetArgSerLeuGlyPhePheSerThrAlaAspSerLeuValAlaAsn 192
QY 316 AATTGGTGGAGCGCTTAGTCAGAGGTGAGGTTCCAGAGCTCGCTGTTTCTATGGC 375
Db 193 AsnLeuValLeuAlaIleTyrArgHisIleThrAsnProGluCysArgGlnTyrIleLeu 212
QY 376 TTTCAAATCTCATCGAATCTTCACTCAGAGATGTACAGTTTGTCTGATAGACTTAC 435
Db 213 ArgGlnAlaPheGluGluAlaIleHisThrHisAlaTyrGlnTyrCysIleGluSerLeu 232
QY 436 ATCAGAGTCCCAAGAAAGGGAATTTTATTTAATGCAATTAAGAACCACTGCTATGTT 495
Db 233 GlyMetAspGluGlyGlu-----ValPheAsnMetTyrArgGluIleProSerVal 249

QY 496 AAGAAAAAAGCAGATTGGCGCTTGCATGG-----ATAGCAGATAGAAAAATCTACT 546
Db 250 AlaArgGlyAlaAlaTrpGlyLeuLysTyrThrGlnSerLeuSerAspProThrPheHis 269
QY 547 TTTGGG-----GAAAGATGGTGGCTTTTGGCTGTGTA 579
Db 270 ThrGlyThrProGluAsnAspGlnArgLeuLeuArgAsnLeuIleAlaPheTyrCysVal 289
QY 580 ---GAAGAGATTTCTTCTCAGGATCTTTTGGCTGTATATCTGGCTTAAGAAAGAGGTT 636
Db 290 LeuGluGlyIlePhePheTyrCysGlyPheThrGlnIleLeuSerMetGlyArgAsn 309
QY 637 CTTATCCAGGACTCCTCTTTTCCAATGAACACTCAGCAGAGATGAAGGACTTCACTGT 696
Db 310 LysMetAsnGlyValAlaGluGlnPheGlnTyrIleLeuArgAspGluSerMetHisLeu 329
QY 697 GACTTTGTGCTGCTGATGTTTC---CAATACTTGAATAAATAGCCTTCA----- 741
Db 330 AsnPheGlyIleAspMetIleAsnGlnIleLysIleGluAsnProHisLeuTrpThrAla 349
QY 742 -----GAAGAAAGGCTCAGGAGATCATTTGTCATGCTCTCAAAATGACGAGGTTT 795
Db 350 GluPheGlnGlnGluValIleGlnMetIleLeuGluGlyThrMetLeuGluIleGluTyr 369
QY 796 TTAACAGAAAGCCTTGCAGTTGGCCTCATTTGGATGAATTCATTTTGTAGAAACAGTAC 855
Db 370 AlaArgAspThrMetProArgGlyValLeuGlyMetAsnAlaSerMetMetGluGluTyr 389
QY 856 ATTGAGTTTGTAGCTGACAGATTAATTTGTTGGAATCTTCAAAAGTTTTCAG--- 912
Db 390 LeuLysPheIleCysAsnArgArgLeuSerGlnLeuGlyLeuProGluGlnPheAlaGly 409
QY 913 GCAGAAAAATCCTTTTGAATTTATG---GAAACATTTCTTTAGAGGAAAAAATTTTC 969
Db 410 ValThrAsnProPheAlaTrpMetSerGluMetMetAspLeuArgLysGluLysAsnPhe 429
QY 970 TTTGAGAAACGAGTTTTCAGAGTATCAG 996
Db 430 PheGluThrArgValThrAspTyrGln 438

RESULT 14

US-09-252-991A-19358
; Sequence 19358, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19358
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19358

Alignment Scores:
Pred. No.: 2,32e-36 Length: 476
Score: 377.50 Matches: 100
Percent Similarity: 45.87% Conservative: 72
Best Local Similarity: 26.67% Mismatches: 154
Query Match: 20.45% Indels: 49
DB: 4 Gaps: 11

US-10-698-228-2 (1-1053) x US-09-252-991A-19358 (1-476)

QY 4 GCGCACCGGAAAGCGGAGCGCGGCTGATCAGGATGAG-----AGATCA 54

```
Db 99 GlySerValGluGluAlaAlaAlaValSerAlaAspSerAlaValAlaAlaAla 118
QY 55 TCTTCAGACACCAAGTAAAGTAAAGTCAATCAAGAGCCACTCTCAAGAAAGAGT 114
Db 119 LysLysAlaLeuAsnAspLeuAspIleGlnGluGlyLeuAspAspLeuGlySerAla 138
QY 115 TCTCGCGGCTTTGTC-----ATC 132
Db 139 AlaArgValGlnValGlyAspLysGlnMetIleAsnAlaArgAlaAspLeuAsnGlnLeu 158
QY 133 TTTCCAAATCAGTACCTGATTTTGGAAAATGATATAACAGCACAGGCTCTCTCTGG 192
Db 159 ValProPheLysTyrAspTrpAlaTrpGlnLysTyrLeuAspGlyCysAlaAsnHisTrp 178
QY 193 ACAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCAGCTGGAAACAG-----CTT 243
Db 179 MetProGlnGluValAsnMetAsnAlaAspIleAlaLeuTrpLysSerLysAspGlyLeu 198
QY 244 AAAGCAGATGAGAGTACTCTCTCATCTTACCTTAGCGCTTTTTCAGCCAGTGATGA 303
Db 199 SerGluHisGluArgIleValMetArgAsnLeuGlyPheSerThrAlaAspSer 218
QY 304 ATTGTAATGAAATTTGGTGAGCGCTTTTAGTCAGGAGTGCAGGTTCCAGAGGCTCGC 363
Db 219 LeuValAlaAsnAsnLeuValLeuAlaValTyrArgLeuIleThrAsnProGluCysArg 238
QY 364 TGTTCCTATGGCTTCAAAATTCATCGAGATGTTCTCAGAGATGTTACAGATGTTGCTG 423
Db 239 GlnTyrIleLeuArgGlnAlaPheGluGluAlaIleHisThrHisAlaTyrGlnTyrCys 258
QY 424 ATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTGAACC 483
Db 259 IleGluSerLeuGlyMetAspGluGlyGlu-----IlePheAsnMetTyrHisGlu 275
QY 484 ATGCCCTATGTTAAGAAAAAGCAGATTGGCGCTTCGGATGGATA----- 528
Db 276 IleProSerValAlaLysLysAlaSerTrpGlyLeuLysTyrThrArgSerIleSerAsp 295
QY 529 -----CGAGATAGAAATCTACTTTTGGGGAAGAGTG 561
Db 296 ProMetPheGlnThrGlyThrProGluThrAspArgGln-----PheLeuArgAsnLeu 313
QY 562 GTGGCTTTGCTGCTGTA---GAAGAGTTTCTTCTCAGGATCTTTTGTGCTATATTC 618
Db 314 IleAlaTyrTyrCysValLeuGluGlyIlePhePheTyrCysGlyPheThrGlnIleLeu 333
QY 619 TGGCTAAAGAAAGAGGCTTATGCGCAGGACTCAGCTTTTCCAACTCACTCAGCAGA 678
Db 334 SerMetGlyArgArgAsnLysMetThrGlyThrAlaGluGlnPheGlnTyrIleLeuArg 353
QY 679 GATGAAGAGCTTCAGTGTGACTTTGCTTGC---CTGATGTTCCAACTAGTAAATAAG 735
Db 354 AspGluSerMetHisLeuAsnPheGlyIleAspValIleAsnGlnIleLysIleGluAsn 373
QY 736 CTT-----TCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTC 777
Db 374 ProHisLeuTrpAspAlaGlnMetLysAspGluAlaThrGlnMetIleLeuGlnGlyThr 393
QY 778 AAAATTTGACGAGGTTTTTAACAGAGCCTTGCCAGTTGGCTCATTTGGAATGATTCG 837
Db 394 GlnLeuGluIleGluTyrAlaArgAspThrMetProArgGlyValLeuGlyMetAsnAla 413
QY 838 ATTTTGTGAAACAGTACATTGAGTTGTAGCTGACAGATTACTTTGTGGAATCTGGATTC 897
Db 414 AlaMetMetGluAspTyrLeuLysPheIleAlaAsnArgArgLeuThrGlnIleGlyLeu 433
QY 898 TCAAGAGTTTTCAG---GCAGAAATCCTTTTGTATTTATGGAACAACT---TCTTTA 951
Db 434 LysGluGluTyrProGlyThrThrAsnProPheProTrpMetSerGluIleMetAspLeu 453
QY 952 GAAGGAAACAAATTTCTTCAGAAACGAGTTTCAGAGTATCAG 996
```

Db 454 LysLysGluLysAsnPheGluThrArgValIleGluTyrGln 468

RESULT 15

US-09-198-452A-1059
; Sequence 1059, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffois, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1059
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-1059

Alignment Scores:
Pred. No.: 2,78e-31 Length: 365
Score: 335.00 Matches: 84
Percent Similarity: 46.84% Conservative: 64
Best Local Similarity: 26.58% Mismatches: 134
Query Match: 18.15% Indels: 34
Gaps: 6
DB:

US-10-698-228-2 (1-1053) x US-09-198-452A-1059 (1-365)

```
QY 130 ATCTTTCCATCCAGTACCTGATATTTGGAAAATGTATATAACAGGCACAGCTTCCTTC 189
Db 50 LeuValProIleLysTyrLysTrpAlaTrpHisTyrLeuAsnGlyCysAlaAsnAsn 69
QY 190 TGCACAGCAGAGAGGTGCGACTTATCAAGGATCTCCTCTCACTGG-----AACAGGTT 243
Db 70 TrpLeuProThrGluValProMetAlaArgAspIleGluLeuTrpLysSerAspGluLeu 89
QY 244 AAAGCAGATGAGAGTACTCTCATCTCTCACATCTTAGCTTTTTCAGCCAGCTGATGA 303
Db 90 SerGluAspGluArgValIleLeuLeuAsnLeuGlyPhePheSerThrAlaGluSer 109
QY 304 ATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGGTTCAGAGGCTCGC 363
Db 110 LeuValGlyAsnAsnIleValLeuAlaIlePheLysHisIleThrAsnProGluAlaArg 129
QY 364 TGTTCCTATGCTTCAATTTCTCATCGAATCTTCTCAGTAAATGTTTATTTAATGCAATTCGAACC 483
Db 130 GlnTyrLeuLeuArgGlnAlaPheGluGluAlaValHisThrHisThrPheLeuTyrIle 149
QY 424 ATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTCGAACC 483
Db 150 CysGluSerLeuGlyLeuAspGluGlyGlu-----ValPheAsnAlaTyrAsnGlu 166
QY 484 ATGCCCTATGTTAAGAAAAAGCAGATTGGCGCTTCGATGATGATGAGAT----- 534
Db 167 ArgAlaSerIleArgAlaLysAspPheGlnMetThrLeuThrValAspValLeuAsp 186
QY 535 -----AGAAATCTACTTTTGGG 552
Db 187 ProAsnPheSerValGlnSerSerGluGlyLeuGlnPheIleLysAsnLeuValGlu 206
QY 553 GAAAGAGTGTGGCTTTGCTGTGTAGAGAGGAGTTTCTCTCAGGATCTTTTCTGCTGCT 612
Db 207 TyrTyrIleIle-----MetGluGlyIlePhePheTyrSerGlyPheValMet 222
QY 613 ATATTCTGGCTAAAGAGAGAGGTCTTATGCGAGGACTCACTTTTCCAAATGAACTATC 672
Db 223 IleLeuSerPheHisArgGlnAsnLysMetThrGlyIleGlyGluGlnTyrGlnTyrIle 242
QY 673 AGCAGAGATGAAGGACTTCACTGCTGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 732
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 30, 2005, 06:34:06 ; Search time 65.75 Seconds
(without alignments)
13390.093 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1846

Sequence: 1 atgggcaccgcggaagcc.....tcaccttgatgcagatttt 1053

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1865214 seqs, 418043040 residues

Total number of hits satisfying chosen parameters: 3730428

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10698228/runat_26102005_100610_4101/app_query.fasta_1.2446
-DB=Published Applications AA -QMTX=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10698228@cgn2_1_113@runat_26102005_100610_4101
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOC=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp.*
- 2: /cgn2_6/ptodata/1/pubpaa/ECT_NEW_PUB.pcp.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pcp.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pcp.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pcp.*
- 6: /cgn2_6/ptodata/1/pubpaa/ECTUS_PUBCOMB.pcp.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pcp.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pcp.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pcp.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pcp.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pcp.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pcp.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pcp.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pcp.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10F_NEW_PUB.pcp.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pcp.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pcp.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pcp.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	98.6	351	15	US-10-698-228-1
2	1454	78.8	389	15	US-10-698-228-5
3	1454	78.8	389	16	US-10-408-765A-559
4	1454	78.8	389	16	US-10-733-878-457
5	1454	78.8	389	18	US-10-756-149-4899
6	1454	78.8	413	9	US-09-925-301-1347
7	1454	78.8	453	18	US-10-220-335-686
8	1454	78.8	453	18	US-10-450-763-45646
9	1425.5	77.2	386	15	US-10-403-571-76
10	1348	73.0	393	20	US-11-097-143-19182
11	1242	67.3	430	15	US-10-369-493-4122
12	1230	66.6	405	14	US-10-128-714-8245
13	1214	65.8	381	15	US-10-369-493-5686
14	1192.5	64.6	381	14	US-10-128-714-3245
15	1192	64.6	413	14	US-10-032-585-7537
16	1176.5	63.7	321	15	US-10-369-493-10365
17	1169.5	63.4	399	15	US-10-369-493-22170
18	1161.5	62.9	391	15	US-10-369-493-22565
19	1161.5	62.9	401	15	US-10-320-797-3145
20	1147.5	62.2	329	15	US-10-424-599-215866
21	1133.5	61.4	717	16	US-10-425-115-215845
22	1132.5	61.3	345	16	US-10-437-963-113592
23	1128.5	61.1	339	16	US-10-767-701-45287
24	1125.5	61.0	339	16	US-10-425-115-247984
25	1114.5	60.4	329	15	US-10-424-599-269685
26	1097.5	59.5	339	15	US-10-424-599-183313
27	1092.5	59.2	400	16	US-10-425-115-200125
28	1061	57.5	347	16	US-10-739-930-6922
29	916	49.6	270	18	US-10-220-335-342
30	871.5	47.2	344	15	US-10-369-493-13401
31	840	45.5	254	16	US-10-425-115-215843
32	805	43.6	345	15	US-10-369-493-21987
33	784	42.5	225	16	US-10-425-115-215840
34	701	38.0	138	15	US-10-264-237-2195
35	680	36.8	335	18	US-10-450-763-45645
36	662	35.9	181	18	US-10-450-763-37219
37	564.5	30.6	338	14	US-10-156-761-10564
38	544	29.5	156	15	US-10-424-599-215870
39	491	26.6	307	15	US-10-369-493-19363
40	456.5	24.7	145	16	US-10-767-701-45286
41	397	21.5	398	15	US-10-282-122A-51357
42	384.5	20.8	415	15	US-10-282-122A-66252
43	384	20.8	403	15	US-10-282-122A-50118
44	383.5	20.8	415	15	US-10-282-122A-69814
45	380.5	20.6	416	15	US-10-282-122A-68043

ALIGNMENTS

RESULT 1

US-10-698-228-1
; Sequence 1, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/698, 228
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019, 733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 1


```
QY 391 GAGATGTTCTCAGATGATACAGTTTCTGATAGACACATTTACATCAGAGATCCAAAG 450
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAAGGGAATTTTATTAATGAATGAACCATGCCCTATGTTTAAAGAAAAGCAGAT 510
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
QY 511 TGGGCTCTCGATGATAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 570
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaIlePhe 228
QY 571 GTCGCTGTAGAAGAGTTTCTTCTCAGAGTCTTTTCTGCTATATCTGCTTAAAGAAAG 630
Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGGCTTTATGCCAGGACTCATTCTTTTCAATGAATCAATCAATCAATCAATCAATCA 690
Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY 691 CACTGTGACTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 750
Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGAGATCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 810
Db 289 ValArgGluIleIleIleAsnAlaValArgIleGluGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCCTCATTTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 870
Db 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
QY 871 GACAGATTACTTGTGGAATCTGATCTCAAAAGTTTTCAGGAGTCTTTCAGGAGAAATCT 930
Db 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY 931 TTTATGAAAACATTTCTTTAGAAGAAAACAAATTTCTTTGAGAAAACGAGTTTCAGAG 990
Db 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY 991 TATCAGCGTTTTCAGTTATGCGAATGCAATTTGATGAAACAGATGATGATGATGATGAT 1050
Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY 1051 TTT 1053
Db 389 Phe 389
```

RESULT 3

US-10-408-765A-559

; Sequence 559, Application US/10408765A

; Publication No. US20040101874A1

; GENERAL INFORMATION:

; APPLICANT: Ghosh, Soumitra S.

; APPLICANT: Fahy, Eoin D.

; APPLICANT: Zhang, Bing

; APPLICANT: Gibson, Bradford W.

; APPLICANT: Taylor, Steven W.

; APPLICANT: Glenn, Gary M.

; APPLICANT: Warnock, Dale E.

; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION

; FILE REFERENCE: 660088.465

; CURRENT APPLICATION NUMBER: US/10/408,765A

; CURRENT FILING DATE: 2003-04-04

; NUMBER OF SEQ ID NOS: 3077

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 559

; LENGTH: 389

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-408-765A-559

Alignment Scores:

Pred. No.: 9.03e-148 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 16 Gaps: 0

US-10-698-228-2 (1-1053) x US-10-408-765A-559 (1-389)

```
QY 91 GAAGAGCCTCTCTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTCCATCCAGTACCCT 150
Db 69 AspGluProLeuLeuArgGluAsnProArgArgPheValIlePheProIleGluTyrHis 88
QY 151 GATATTGGAAATGTATTAACAGGCACAGCTTCTCTGACAGCAGAGAGAGCTCCAC 210
Db 89 AspIleTrpGlnMetTyrLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
QY 211 TTATCAAGAGATCTCCCTCTCACTGGAACAAGCTTTAAAGCAGATGAGAAGTACTTCTCTCT 270
Db 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTTAGCCTTTTTCAGCCAGTGATGGAATTTGTAATGAAATTTTGGTGGAGCGC 330
Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY 331 TTTAGTCAGGAGTGCAGGTTCCAGAGCTCGCTGTTTCTATGCTTTCATGCTTCAATTTCT 390
Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGAATGTTTCACTCAGAGATGTACAGTTTGTGTATAGACACTTACATCAGAGATCCCAAG 450
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAAGGGAATTTTATTTAATGCAATTTGAAACCACTGCTTATGTTAAAGAAAACAGAT 510
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
QY 511 TGGGCTTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 570
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaIlePhe 228
QY 571 GTCGCTGTAGAAGAGTTTCTTCTCAGAGTCTTTTCTGCTATATCTGCTTAAAGAAAG 630
Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGGCTTTATGCCAGGACTCATTCTTCCAAATGCAATCTCATCAGCAGATGAAAGACTT 690
Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY 691 CACTGTGACTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 750
Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGAGATCATTTGATGCTGCTCAAAATTCAGCAGGAGTCTTTTAAACAGAGCTTTG 810
Db 289 ValArgGluIleIleIleAsnAlaValArgIleGluGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCCTCATTTGAATGCAATTTGATGAAACAGATGATGATGATGATGATGATGAT 870
Db 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
QY 871 GACAGATTACTTGTGGAATCTGATCTCAAAAGTTTTCAGGAGTCTTTCAGGAGAAATCT 930
Db 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY 931 TTTATGAAAACATTTCTTTAGAAGAAAACAAATTTCTTTGAGAAAACGAGTTTCAGAG 990
Db 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY 991 TATCAGCGTTTTCAGTTATGCGAATGCAATTTGATGAAACAGATGATGATGATGATGAT 1050
Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
```



```
QY 1051 TTT 1053
Db 389 Phe 389

RESULT 4
US-10-733-878-457
; Sequence 457, Application US/10733878
; Publication No. US20040224408A1
; GENERAL INFORMATION:
; APPLICANT: Jean-Philippe Girard
; APPLICANT: Francois Amalric
; APPLICANT: Myriam Roussigne
; APPLICANT: Thomas Clouaire
; TITLE OF INVENTION: THAP PROTEINS AS NUCLEAR RECEPTORS FOR
; TITLE OF INVENTION: CHEMOKINES AND ROLES IN TRANSCRIPTIONAL REGULATION, CELL
; TITLE OF INVENTION: PROLIFERATION AND CELL DIFFERENTIATION
; FILE REFERENCE: BIOBANK.012A
; CURRENT APPLICATION NUMBER: US/10/733,878
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 60/432699
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/485027
; PRIOR FILING DATE: 2003-07-03
; NUMBER OF SEQ ID NOS: 535
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 457
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-733-878-457

Alignment Scores:
Pred. No.: 9.03e-148 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 16 Gaps: 0

US-10-698-228-2 (1-1053) x US-10-733-878-457 (1-389)

QY 91 GAAGGCCACTCTTAAGAAAGAGTTCTCGCGGTTGTGTCATCTTTCCAATCCAGTACCCT 150
Db 69 AspGluProLeuLeuArgGluAsnProArgArgPheValIlePheProIleGluTyrHis 88

QY 151 GATATTGGAAATGTATTAACAGGCACAGGCTTCTCTGGACAGCAGAGAGGTCGAC 210
Db 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108

QY 211 TTATCAAGGATCTCCCTCACTGGAACAAGCTTAAGCAGATGAGAGTACTCTCTCT 270
Db 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128

QY 271 CACATCTTAGCCTTTTTCAGCCAGTGTGAATGTAATGAAATGAAATTTGGTGGAGCGC 330
Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148

QY 331 TTTAGTCAGAGGTGCGAGTTCAGAGGCTCGCTGTCTTCTATGGCTTTCAAAATCTCATC 390
Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168

QY 391 GAGATGTTCACTCAGAGATGTACAGTTTGGTGTATGACACTTACATCAGAGATCCCAAG 450
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188

QY 451 AAAAGGGAATTTTATTTAATCAATGAAACCATGCCCTATCTTTAAGAAAAAGCAGAT 510
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208

QY 511 TGGGCTTCGGATGAGATAGATGAAATCTACTTTTGGGAAAGAGTGGTGGCCTTT 570
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValValAlaPhe 228

571 GCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTCTGCTCTATATTCTGCTAAAGAAG 630
Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248

QY 631 AGAGTCTTTATGCCAGGACTCACTTTTTCAAATCAATCAGCAGAGATGAAGGACTT 690
Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268

QY 691 CACTGTGACTTTTCTGCTGATGTTCCCAATACCTAGTAATAAGCCTTCAGAGAAGG 750
Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288

QY 751 GTCAGGAGATCATTTGTGATGCTGCAAAATTTGAGCAGGAGTTTAAACAGAGCCCTTG 810
Db 289 ValArgGluIleIleIleAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308

QY 811 CCAGTTGGCCTCATTCGAATGCAATTTGATGAAACAGTACATGAGTTGTAGCT 870
Db 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328

QY 871 GACAGATTACTTGTGAACCTTGAATTTCAAGGTTTTCAGGCAGAGAAATCCCTTTGAT 930
Db 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348

QY 931 TTTATCGAAAAACATTTCTTTAGAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAGAG 990
Db 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368

QY 991 TATCAGCGTTTTCAGTTATGGCAGAAACCCAGATAACCTCTTTCACCTTGGATGCAGAT 1050
Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388

QY 1051 TTT 1053
Db 389 Phe 389

RESULT 5
US-10-756-149-4899
; Sequence 4899, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4899
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-756-149-4899

Alignment Scores:
Pred. No.: 9.03e-148 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 18 Gaps: 0

US-10-698-228-2 (1-1053) x US-10-756-149-4899 (1-389)

QY 91 GAAGGCCACTCTTAAGAAAGAGTTCTCGCGGTTGTGTCATCTTTCCAATCCAGTACCCT 150
Db 69 AspGluProLeuLeuArgGluAsnProArgArgPheValIlePheProIleGluTyrHis 88

QY 151 GATATTGGAAATGTATTAACAGGCACAGGCTTCTCTGGACAGCAGAGAGGTCGAC 210
Db 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108

QY 211 TTATCAAGGATCTCCCTCACTGGAACAAGCTTAAGCAGATGAGAGTACTCTCTCT 270
Db 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128

QY 271 CACATCTTAGCCTTTTTCAGCCAGTGTGAATGTAATGAAATGAAATTTGGTGGAGCGC 330
Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148

QY 331 TTTAGTCAGAGGTGCGAGTTCAGAGGCTCGCTGTCTTCTATGGCTTTCAAAATCTCATC 390
Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168

QY 391 GAGATGTTCACTCAGAGATGTACAGTTTGGTGTATGACACTTACATCAGAGATCCCAAG 450
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188

QY 451 AAAAGGGAATTTTATTTAATCAATGAAACCATGCCCTATCTTTAAGAAAAAGCAGAT 510
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208

QY 511 TGGGCTTCGGATGAGATAGATGAAATCTACTTTTGGGAAAGAGTGGTGGCCTTT 570
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValValAlaPhe 228
```

Db 89 AspileTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
QY 211 TTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAGTACTTCTCTCT 270
Db 109 LeuSerLysAspileGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTTAGCCCTTTTTCAGCCAGTGTGAATTTGTAATGAAATTTGGTGAGCGC 330
Db 129 HisValLeuAlaPhePheAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY 331 TTATGTCAGGAGTGCAGGCTTCAGAGGCTCGCTGTTTCTATGGCTTTCAAAATCTCATC 390
Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGATGTTCACTCAGAGATGTACAGTGTGCTGATAGACACTTACATCAGAGATCCCAAG 450
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAAGGGAATTTTATTAAATGCAATTTGAAACCATGCCCTATGTTTAAAGAAAAAGCAGAT 510
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
QY 511 TGGGCTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 570
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 228
QY 571 GCTGCTGTAGAAGGATTTTCTCTCAGAGATCTTTGCTGCTATATCTTGGCTAAAGAAG 630
Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGGTCTTATGCCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGACTT 690
Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY 691 CACTGTGACTTTGCTGCTGATGTTTCCAAATGAACTTAAAGGCTTTCAGAGAAAGG 750
Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGAGATCATTTGCTGATGCTGCAAAATTTGACAGAGGTTTAAACAGAGCTTG 810
Db 289 ValArgGluIleIleAlaValArgIleGluGluGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCTCATTTGGAATGAATTCATTTGATGAAACAGTACATTCAGTTGTAGCT 870
Db 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
QY 871 GACAGATTAATTTGTAATGCAATTTGAAACCAATGCCCTATGTTTAAAGAAAAAGCAGAT 930
Db 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY 931 TTTATGGAACATTTCTTTAGAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAGAG 990
Db 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY 991 TATCAGGCTTTGTCAGTTATGTCAGAAACCAAGATTAAGCTTTCACCTGATGAGAT 1050
Db 369 TyrGlnArgMetGlyValMetSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY 1051 TTT 1053
Db 389 Phe 389

RESULT 6

US-09-925-301-1347
; Sequence 1347, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882

; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1347
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1347

Alignment Scores:
Pred. No.: 9,23e-148 Length: 413
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: Gaps: 0

US-10-698-228-2 (1-1053) x US-09-925-301-1347 (1-413)

QY 91 GAAGAGCCACTCTCTAAGAAAGAGTTCTGCGGGTTTGTCTCATCTTTCATCCAGTACCT 150
Db 93 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 112
QY 151 GATATTTGGAATGTATAACAGGACAGGCTTCTCTCGACAGCAGAGAGGTGAC 210
Db 113 AspileTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 132
QY 211 TTATCAAGGATCTCCCTCACTGCAACAAGCTTAAAGCAGATGAGAGTACTTCTCTCT 270
Db 133 LeuSerLysAspileGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 152
QY 271 CACATCTTAGCCCTTTTTCAGCCAGTGTGAATTTGAAATTTGGTGAGCGC 330
Db 153 HisValLeuAlaPhePheAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 172
QY 331 TTATGTCAGGAGTGCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAAATTTCTCATC 390
Db 173 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 192
QY 391 GAGATGTTCACTCAGAGATGTACAGTTTCTGATGACACTTACATCAGAGATCCCAAG 450
Db 193 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 212
QY 451 AAAAGGGAATTTTATTAAATGCAATTTGAAACCAATGCCCTATGTTTAAAGAAAAAGCAGAT 510
Db 213 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 232
QY 511 TGGGCTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 570
Db 233 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 252
QY 571 GCTGCTGTAGAAGGATTTTCTCTCAGAGATCTTTGCTGCTATATCTGCTATAATCTGCTAAAGAG 630
Db 253 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 272
QY 631 AGAGGTCTTATGTCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGGACTT 690
Db 273 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 292
QY 691 CACTGTGACTTTGCTGCTGATGTTTCCAAATGAACTTAGTAAATAGGCTTTCAGAGAAAGG 750
Db 293 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 312
QY 751 CTCAGGAGATCATTTGATGCTGTCAAAATTTGAGCAGGAGTGTAAACAGAGAGCTTG 810
Db 313 ValArgGluIleIleAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 332
QY 811 CCAGTTGGCTCATTTGGAATGAATTTGATGAAACAGTACATTCAGTTGTAGCT 870
Db 333 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 352

QY 871 GACGATTACTTGTGAAGTCTGCAATCTCAAGGTTTTCAGGCAGAAATCCCTTTTGAT 930
Db 353 AspArgLeuMetLeuGluGlyPheSerLysValPheArgValGluAsnProPheAsp 372
QY 931 TTTATGGAACATTTCTTAGAGGAAACAAATTTCTTGAGAAACGAGTTTCAGAG 990
Db 373 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 392
QY 991 TATCAGCGTTTTCAGTCTTATGCGAGAAACACAGATAAAGCTCTTACCTTGGATGCAGAT 1050
Db 393 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 412
QY 1051 TTT 1053
Db 413 Phe 413
RESULT 7
US-10-220-335-686
; Sequence 686, Application US/10220335
; Publication No. US20050175607A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-041
; CURRENT APPLICATION NUMBER: US/10/220,335
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 09/664,641
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,807
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/597,707
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/577,409
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 09/515,126
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 688
; SOFTWARE: Custom
; SEQ ID NO 686
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-220-335-686
Alignment Scores:
Pred. No.: 9.56e-148 Length: 453
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 18 Gaps: 0
US-10-698-228-2 (1-1053) x US-10-220-335-686 (1-453)
QY 91 GAAGAGCCACTCTTAAGAAAGTCTCGCGGTTTGTCTATCTTCCAAATCCAGTACCCT 150
Db 133 AspGluProLeuMetLeuGluAsnProArgPheValIlePheProIleGluThrHis 152
QY 151 GATATTGGAATAATGATAACAGGCACAGGCTCTCTTCTGACAGCAGAAAGGTCGAC 210
Db 153 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 172
QY 211 TATCAAAAGATCTCCCTCTCTGGAACACAGCTTAAAGCAGATGAGAAGTACTCTCTCT 270
Db 173 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 192
QY 271 CACATCTTAGCCTTTTTCAGCCAGTGTGAATTTGTAATGAAATTTGGTGGAGCGC 330
Db 193 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 212
QY 331 TTTAGTCAGGAGTGCAGGTTCCAGAGGCTCGCTGTTCTTATGGCTTTCAAATTTCTCATC 390
Db 213 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 232

QY 391 GAGAATGTTCACTCAGAGATGTACAGTTGCTGATACACACTTACATCAGAGATCCCAAG 450
Db 233 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 252
QY 451 AAAAGGAAATTTTATTAATCAATGAAACCATGCCCTATGTTAAGAAAAAGCAGAT 510
Db 253 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAsp 272
QY 511 TGGGCTTGGGATGGATAGCAGATAGAAATCTACTTTTGGGAAAGAGTGGTGGCTTT 570
Db 273 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 292
QY 571 GCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTGCTGCTATATTTCTGGCTAAAGAG 630
Db 293 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 312
QY 631 AGAGTCTTATGCCAGGACTCATTTTCCAAATCAATCAATCAGCAGAGATGAAGGCTT 690
Db 313 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 332
QY 691 CACTGTGACTTTCCTTGGCTGATGTTCCCAATCTAGTAATTAAGCCTTCAGAAAGAGG 750
Db 333 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 352
QY 751 GTCAGGAGATCATTTGTTGATGCTGTCAAAATTTGAGCAGGAGTTTAAACAGAGCCTTG 810
Db 353 ValArgGluIleIleIleAsnAlaValArgIleGluGluPheLeuThrGluAlaLeu 372
QY 811 CCAGTTGGCTCTCATTTGGAATGAAATTTGATGAACAGTACATTTGAGTTTGTAGCT 870
Db 373 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 392
QY 871 GACAGATTACTTGTGAACTTCTCAAGTTTTCAGGAGGAGAAATCCCTTTGAT 930
Db 393 AspArgLeuMetLeuGluGlyPheSerLysValPheArgValGluAsnProPheAsp 412
QY 931 TTTATGAAAAACATTTCTTTAGAAAGGAAAAACAAATTTCTTTCAGAAACAGAGTTTCAGAG 990
Db 413 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 432
QY 991 TATCAGCGTTTTCAGTCTTATGCGAGAAACACAGATAAAGCTCTTTCACCTTGGATGCAGAT 1050
Db 433 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 452
QY 1051 TTT 1053
Db 453 Phe 453
RESULT 8
US-10-450-763-45646
; Sequence 45646, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 45646
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN

; LOCATION: (148)..(203)
; OTHER INFORMATION: Ribonucleotide reductase small subunit proteins domain
; OTHER INFORMATION: identified by eMATRIX, accession number BL00368A, p-value=1.000e-
; OTHER INFORMATION: 40, raw score of 36.98
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (167)..(440)
; OTHER INFORMATION: Ribonucleotide reductases domain identified by Pfam,
; OTHER INFORMATION: accession name ribonuc_red, E-value=9e-174, Pfam score of 507.8
US-10-450-763-45646

Alignment Scores:
Pred. No.: 9,56e-148 Length: 453
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservativity: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 18 Gaps: 0

US-10-698-228-2 (1-1053) x US-10-450-763-45646 (1-453)

QY	91	GAAGAGCCACTCTCTAGAAAGAGTTCTCGCGGTTTGTCTATCTTCCATCCAGTACCCT	150
DB	133	AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis	152
QY	151	GATATTGGAAATGTATAACAGGCACAGGCTTCTCTCGACAGCAGAGAGGTCGAC	210
DB	153	AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluValAsp	172
QY	211	TTATCAAGAGATCCCTCAGTGAACAAAGCTTAAAGCAGATGAGAGTACTTCTCT	270
DB	173	LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer	192
QY	271	CACATCTTAGCCTTTTTCAGCCAGTGTGAAATGTAATCAAAATTTGGTGGAGCGC	330
DB	193	HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg	212
QY	331	TTTAGTCAGAGGTGCAGGTTCAGAGGCTCGCTGTTCTCTATGGCTTTTCAAAATCTCATC	390
DB	213	PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet	232
QY	391	GAGATGTTCACTCAGAGATGTACAGTTTCTGCTGATAGACACTTACATCAGAGATCCCAAG	450
DB	233	GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys	252
QY	451	AAAAGGAATTTTATTAATGCAATGAAACCATGCCCTATGTTAAGAAAAAAGCAGAT	510
DB	253	GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAsp	272
QY	511	TGGGCGCTTGGATAGCAGATAGAAATCTACTTTTGGGAAAGAGTGGGCTTTT	570
DB	273	TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaAlaPhe	292
QY	571	GCTGCTGTAGAAGAGTCTTCTCTCAGGATCTTTTGTCTATATCTTGGCTAAAGAAG	630
DB	293	AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys	312
QY	631	AGAGTCTTATGCCAGGATCACTTTTCCAATGAACTCATCAGCAGAGATGAAGACTT	690
DB	313	ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu	332
QY	691	CACTGTGACTTTCCTTGGCTGATGTTTCCAAATCTAGTAATTAAGCTTCCAGAAAGG	750
DB	333	HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg	352
QY	751	GTGAGGAGATCAATGTTGATGCTGTGCAAAATGAGCAGAGTGTTTTAAACAGAGCTTG	810
DB	353	ValArgGluIleIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu	372
QY	811	CCAGTTGGCCTCATTTGGAATGAATTCATTTTGATGAAACAGTACATTTGAGTTTGTAGCT	870
DB	373	ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla	932

QY	871	GACAGATTACTTGTGGAACTTGGATTCTCAAAGGTTTTTTCAGGCAGAGAAATCCTTTGAT	930
DB	393	AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp	412
QY	931	TTTATGGAACATTTCTTTAGAGGAAAAACAATTTCTTTGAGAAACGAGTTTCAGAG	990
DB	413	PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu	432
QY	991	TATCAGCGTTTTCAGATTATGGCAGAAACACAGATAACGCTTTCACCTTGGATGCAGAT	1050
DB	433	TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp	452
QY	1051	TTT 1053	
DB	453	Phe 453	

RESULT 9
US-10-403-571-76
; Sequence 76, Application US/10403571
; Publication No. US20040068763A1
; GENERAL INFORMATION:
; APPLICANT: Hopkins, Nancy
; APPLICANT: Golling, Gregory
; APPLICANT: Amsterdam, Adam
; APPLICANT: Sun, Zhaoxia
; TITLE OF INVENTION: Developmental Mutations in Zebrafish
; FILE REFERENCE: 01997/539002
; CURRENT APPLICATION NUMBER: US/10/403,571
; CURRENT FILING DATE: 2003-03-25
; PRIOR APPLICATION NUMBER: US 60/368,760
; PRIOR FILING DATE: 2002-03-29
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Danio rerio
US-10-403-571-76

Alignment Scores:
Pred. No.: 1,09e-144 Length: 386
Score: 1425.50 Matches: 268
Percent Similarity: 89.41% Conservativity: 36
Best Local Similarity: 78.82% Mismatches: 31
Query Match: 77.22% Indels: 5
DB: 15 Gaps: 1

US-10-698-228-2 (1-1053) x US-10-403-571-76 (1-386)

QY	49	AGATCATCTTCAGACACCAAGAAAGTGAATAAGTCA-----AATGAA	93
DB	47	ArgLysIlePheAspGluSerGluGlyGlnSerLysAlaLysLysGlyAlaValGluGlu	66
QY	94	GAGCCTCTCTAAGAAAGAGTCTCGCGGTTTCTCATCTTTTCCAAATCCAGTACCCTGAT	153
DB	67	GluProLeuLeuLysGluAsnProHisArgPheValIlePheProIleGlnTyrHisAsp	86
QY	154	ATTGGAATAATGATAAACAGGCACAGCTTCTTCTGGACAGCAGAGAGTTCGACTTA	213
DB	87	IleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAspLeu	106
QY	214	TCAAAGATCTCCTCTCATCGAACACAGCTTAAACGACATGAGAGTACTTCTCTCTAC	273
DB	107	SerLysAspLeuGlnHisTrpAspSerLeuLysAspGluGluArgTyrPheIleSerHis	126
QY	274	ATCTTAGCCTTTTTCAGCGCAGTGTGGAATTCGTAATAAATTTGCTGGAGCGCTTT	333
DB	127	ValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArgPhe	146
QY	334	AGTCAGAGGTGCAGGTTCAGAGGCTCGCTGTTTCTATGGCTTTTCAAATTTCTCATCGAG	393
DB	147	ThrGlnGluValGlnValThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMetGlu	166

```
QY 394 AATGTTCACTCAGAGATGACAGTGGTTGCTGATAGACACTTACATCAGAGATCCCAAGAAA 453
Db 167 AsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspSerLysGlu 186
QY 454 AGGGAATTTTATTAATGCAATTTGAACCATGCCCTATGTTAAGAAAAACAGAGATTGG 513
Db 187 ArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAspTrp 206
QY 514 GCCTTGCGATGACAGATGAGAAATCTACTTTTGGGGAAGAGTGGTGGCTTTGCT 573
Db 207 AlaLeuAsnTrpIleCysAspSerLeuLeuIleAspThrTyrIleLysAspSerLysGlu 226
QY 574 GCTGTAGAGAGGATTTTCTTCTCAGGATCTTTGCTGCTATATTTCTGGCTAAAGAGAGA 633
Db 227 AlaValGluGlyLePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArg 246
QY 634 GGTCTTATGCCAGGACTCACTTTTCCAAATGACATCATCAGCAGAGATGAAGACTTCAC 693
Db 247 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHis 266
QY 694 TGTGACTTTGCTGCTGATGTTCCAAATCTAGTAAATAGCCCTTCAGAGAAAGGCTC 753
Db 267 CysAspPheAlaCysLeuMetPheLysHisLeuIleAsnLysProSerGluGluThrVal 286
QY 754 AGCGAGATCATTTGATGCTGTCAAAATGACAGGAGTCTTTTACAGAAGCCTTGCCA 813
Db 287 LysLysIleIleMetAsnAlaValGluIleGluGluPheLeuThrAspAlaLeuPro 306
QY 814 GTTGCCCTCATTTGAATGAATTCATTTTGTGATGAACAGTACATTCAGTGGTTGTAGCTGAC 873
Db 307 ValLysLeuIleGlyMetAsnCysAspLeuMetLysGlnTyrIleGluPheValAlaAsp 326
QY 874 AGATTACTGTGGAAGTGGATCTCAAGGTTTTCAGGCGAAGAAATCTTTGATTTT 933
Db 327 ArgLeuLeuLeuGluGlyPheAspLysValTyrArgValGluAsnProPheAspPhe 346
QY 934 ATGGAAGACATTTCTTGAAGGAAACAAATTTCTTGAAGAACGAGTTTCAGAGTAT 993
Db 347 MetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGluTyr 366
QY 994 CAGCCTTTTGCAGTTATGCGAGAAACCAAGATAACGCTTCACCTTGGATGCAATTTT 1053
Db 367 GlnArgMetGlyValMetSerGlyThrThrAspAsnThrPheThrLeuAspAlaAspPhe 386
RESULT 10
US-11-097-143-19182
; Sequence 19182, Application US/11097143
; Publication No. US2005020858A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
```

```
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19182
; LENGTH: 393
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-19182
Alignment Scores:
Pred. No.: 2,67e-136 Length: 393
Score: 1348.00 Matches: 252
Percent Similarity: 85.92% Conservative: 41
Best Local Similarity: 73.90% Mismatches: 48
Query Match: 73.02% Indels: 0
DB: 20 Gaps: 0
US-10-698-228-2 (1-1053) x US-11-097-143-19182 (1-393)
QY 31 GGCTGGATCAGGATCAGAGATCATCTTCAGACACCAACGAAAGTGAATAAAGTCAAT 90
Db 53 GlyIleGlyLysSerAlaAsnSerLeuMetGluLysSerValThrProPheAspProSer 72
QY 91 GAAGAGCCTCTCTAAGAAAGAGTTCGCCGGTTTGTCTATCTTCCATCCAGTACCCT 150
Db 73 LeuGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGlnTyrHis 92
QY 151 GATATTGGAAATGTATAACAGGCAGGCTCTCTTCGACAGCAGAGAGAGTCCGAC 210
Db 93 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrValGluGluValAsp 112
QY 211 TTATCAAAAGATCTCCCTCACTGCAACCAAGCTTAAAGCAGATGAGAAGTACTTCTCT 270
Db 113 LeuSerLysAspLeuThrAspTrpHisArgLeuLysAspAspGluArgHisPheIleSer 132
QY 271 CACATCTTAGCCTTTTTCAGCCAGCTGATGGAATTTGAAATGAAATTTGGTGAGCGC 330
Db 133 HisValLeuAlaPhePheAlaIleAspGlyIleValAsnGluAsnLeuValGluArg 152
QY 331 TTAGTCAGGAGCTGCAGGTTCCAGAGGCTCGCTGTTCTTCTATGGCTTCAATTTCTC 390
Db 153 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 172
QY 391 GAGAATGTTTCACTCAGAGATGTACAGTTTGTGTAGTACACTTACATCAGAGATCCCA 450
Db 173 GluAsnValHisSerGluMetTyrSerValLeuIleAspThrTyrIleArgAspProHis 192
QY 451 AAAAGGAATTTTATTAATGCAATTTGAACCATGCCCTATGTTAAGAAAAACAGCAT 510
Db 193 GlnArgGluTyrLeuPheAsnAlaIleGluThrMetProAlaValLysArgLysAlaAsp 212
QY 511 TGGGCTTGGGATGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCTTT 570
Db 213 TrpAlaLeuSerTrpIleSerSerLysSerAlaAsnPheGlyGluArgIleAlaPhe 232
QY 571 GCTGCTGTAGAAGGATTTTCTCTCAGGATCTTTTGTGCTGTATATTCTGCTAAAGAG 630
Db 233 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 252
QY 631 AGAGTCTTATGCGCAGGACTCCTCTTTTCCAAATGAACTCATCAGCAGAGATGAAGACT 690
Db 253 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 272
QY 691 CACTGTGACTTTGCTGCTGATGTTCCCAATACATTTAGTAATAAGCTTTCAGAGAAAGG 750
Db 273 HisCysAspPheAlaValLeuMetPheGlnHisLeuValGlnArgProLysArgGluArg 292
QY 751 GTCAGGAGAGATCATTTGTTGATGCTCAAAATTCAGCAGGAGTGTATTAACAGAGCCT 810
Db 293 IleIleGluIleIleArgAspAlaValAlaIleGluGlnGluPheLeuThrAspAlaLeu 312
QY 811 CCAGTTGGCCTCATTTGGAATGAATTTGATTTTGAATGAACAGTACATTTAGTGTAGCT 870
Db 313 ProValAsnLeuIleGlyMetAsnCysAspLeuMetSerGlnTyrIleGluPheValAla 332
```

QY 871 GACAGATTACTTGTGAACCTTGGAATCTCAAGAGTTTTCAGGCAGAGAAATCCTTTTGAT 930
 |||||
 DB 333 AspArgLeuValGluLeuGlyValGlyLysIleTyrAsnThrLysAsnProPheAsn 352
 |||||
 QY 931 TTTATGGAAAACATTTCTTAGAGGAAAACAAATTTCTTTCAGAAAACGAGTTTCAGAG 990
 |||||
 DB 353 PheMetGluMetIleSerLeuAspGlyLysThrAsnPhePheGluLysValGlyGlu 372
 |||||
 QY 991 TATCAGCGTTTTCAGTTATGGCAGAAACCCACAGATAACCTCTTACCTTGGATGCAGAT 1050
 |||||
 DB 373 TyrGlnArgMetGlyValValSerAsnProLeuAspAsnValPheThrLeuAspAlaAsp 392
 |||||
 QY 1051 TTT 1053
 |||||
 DB 393 Phe 393

RESULT 11
 US-10-369-493-4122
 ; Sequence 4122, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; PRIOR FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 4122
 ; LENGTH: 430
 ; TYPE: PRT
 ; ORGANISM: Neurospora crassa
 US-10-369-493-4122

Alignment Scores:
 Pred. No.: 8,12e-125 Length: 430
 Score: 1242.00 Matches: 243
 Percent Similarity: 76.03% Conservative: 52
 Best Local Similarity: 62.63% Mismatches: 47
 Query Match: 67.28% Indels: 46
 DB: 15 Gaps: 6

US-10-698-228-2 (1-1053) x US-10-369-493-4122 (1-430)

QY 28 GCCGGCTGGATCAGATGAGATCATCTTCAGACACCAACGAAAGTGAA----- 78
 |||||
 DB 43 AlaGluIleAspAlaGluHisAsnAlaAsnLysLysAlaAlaGluAlaLysLysMetAla 62
 |||||
 QY 79 -----ATAAGTCAATGAA-----GAGCCACTCTTAAGAAAGAGTTCTCGCGGTTT 126
 |||||
 DB 63 ProThrLeuLysProGluAlaAsnGluProLeuLeuThrGluAsnProGlnArgPhe 82
 |||||
 QY 127 GTCATCTTTTCAATCCAGTACCCCTCAT----- 153
 |||||
 DB 83 ValLeuPheProIleLysTyrHisGluValCysHisAlaSerGlnThrPheGlnThrThr 102
 |||||
 QY 154 -----ATTGGAAAATGTATAAA 171
 |||||
 DB 103 ArgLysValValHisGlyLeuIleGlyTyrProHisTyrArgIleTyrGlnMetTyrLys 122
 |||||
 QY 172 CAGGCACAGGCTCTCTTGGACAGCAGAGAGTGCAGTTATCAAGATCTCCCTCAC 231
 |||||
 DB 123 LysAlaGluAlaSerPheThrAlaGluGluIleAspLeuSerLysAspLeuHisAsp 142
 |||||
 QY 232 TGG---AACAAAGCTTAAAGCAGATGAGAAGTACTCTCTCATCTCTAGCCTTTT 288
 |||||

DB 143 TrpAsnAsnArgLeuAsnAspAspGluLysPhePheIleSerHisIleLeuAlaPhePhe 162
 QY 289 GCAGCCAGTCATGAATTTGTAATAATTTGGTGGAGCGCTTTAGTCAGAGGTCGAG 348
 |||||
 DB 163 AlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlyGluValGln 182
 |||||
 QY 349 GTTCAGAGGCTCGCTGCTTTTCATGGCTTTCAAAATTTCTCATCGAGAAATGTTCACTCAGAG 408
 |||||
 DB 183 IleProGluAlaArgCysPheTyrGlyPheGlnIleMetMetGluAsnIleHisSerGlu 202
 |||||
 QY 409 ATGTACAGTTGCTGTAGTACACATTCATCAGAGATCCCAAGAAAGGAAATTTTATTT 468
 |||||
 DB 203 ThrTyrSerLeuLeuIleAspThrTyrIleLysGluProSerGlnArgThrTyrLeuPhe 222
 |||||
 QY 469 AATGCAATTCGAACCATGCGCTATGTTAAGAAAAGCAGATTCGGCTTCGGATGGATA 528
 |||||
 DB 223 AsnAlaIleAspThrIleProCysIleArgLysLysAlaAspTrpAlaLeuArgTrpIle 242
 |||||
 QY 529 GCAGATAGAAAATCTACTTTTGGGGAAGAGTGTGGCTTTGCTGCTGTAGAAGGAGTT 588
 |||||
 DB 243 ThrAspLysSerSerThrPheAlaGlnArgLeuValAlaPheAlaValGluGlyIle 262
 |||||
 QY 589 TTCTTCTCAGGATCTTTTGTGCTATATTCGGCTAAAGAGAGAGGCTTTATGCCAGGA 648
 |||||
 DB 263 PhePheSerGlyAlaPheAlaSerIlePheTrpLeuLysLysArgGlyLeuMetProGly 282
 |||||
 QY 649 CTCACTTTTCCATGATCAATCAGCAGATCAAGGACTTCACTGTGACTTGTGCTTGC 708
 |||||
 DB 283 LeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisThrAspPheAlaCys 302
 |||||
 QY 709 CTGATGTTTCCAACTACTAGTAAATAAGCCTTCAGAAGAAAGGTCAGGAGATCATTTGT 768
 |||||
 DB 303 LeuLeuPheSerHisLeuAsnAsnArgProSerLysGlnLeuIleGlnIleVal 322
 |||||
 QY 769 GATGCTGTCAAAATTCAGCAGGAGCTTTTAAACAGAAGCTTGGCCTCATTTGA 828
 |||||
 DB 323 AspAlaValArgIleGluGlnPheLeuThrGluAlaLeuProCysAlaLeuLeuGly 342
 |||||
 QY 829 ATGAATTGCTTTGTAGTAAACAGTACATTCAGTGTGTTGTAGCTGACAGATTACTTGTGAA 888
 |||||
 DB 343 MetAsnAlaAspLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValAla 362
 |||||
 QY 889 CTTGGATCTCAAGGTTTTCAGGCAGAAAATCCTTTTATGTTTATCGAAAACATTTCT 948
 |||||
 DB 363 LeuGlyAsnGluLysIleTyrArgSerThrAsnProPheAspPheMetGluAsnIleSer 382
 |||||
 QY 949 TTAGAAGGAAAAACAAATTTCTTTGAGAAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTT 1008
 |||||
 DB 383 LeuGlyGlyLysThrAsnPhePheGluLysArgValGlyAspTyrGlnLysAlaGlyVal 402
 |||||
 QY 1009 ATG-----GCAGAAAACCCAGATAACGTC----- 1032
 |||||
 DB 403 MetAsnSerThrLysLysAlaAspAlaGluValAlaLysAsnGluAsnGlyGly 422
 |||||
 QY 1033 ---TTCACCTTGGATCGAGATTTT 1053
 |||||
 DB 423 AspPheThrPheAspGluAspPhe 430
 |||||

RESULT 12

US-10-128-714-8245
 ; Sequence 8245, Application US/10128714
 ; Publication No. US20030119013A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jiang, Bo
 ; APPLICANT: Hu, Weng
 ; APPLICANT: Tishkoff, Daniel
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Eroshkin, Alexey M
 ; APPLICANT: Lemieux, Sebastien M
 ; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
 ; FILE REFERENCE: 10182-018-999
 ; CURRENT APPLICATION NUMBER: US/10/128,714

[illegible]

RESULT 14

```

RESUM 14
US-10-128-714-3245
; Sequence 3245, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and

```

Db 233 ProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisThrAspPhe 252
QY 703 GCTTCCCTGATGCTTCAATAGCTTAAATAGCCTTCAAGAAAGGCTCAGGAGATC 762
Db 253 AlaCysLeuLeuPheSerHisLeuAsnHisArgProAspLysLysValIleGluAspIle 272
QY 763 ATTGTTGATGCTGCTCAAAATTCAGCAGGAGTCTTTAAACAGAAAGCCTTGCCAGTTGCCCTC 822
Db 273 IleValGluAlaValAlaIleGluGlnPheLeuThrAspAlaLeuProValAlaLeu 292
QY 823 ATTGAATGAATTCATTTGATGAACAGTACATGATGCTTGTAGCTCAGATTAATTT 882
Db 293 LeuGlyMetAsnSerLysLeuMetCysGlnIleGluPheValAlaAspArgLeuLeu 312
QY 883 GTGGAACTTGGATCTCAAGAGTCTTTCAGGCAGAAATCTTTGATTTATGGAAAC 942
Db 313 ValAlaLeuGlyAsnLysLysTyrPheAsnSerProAsnProPheAspPheMetGluSer 332
QY 943 ATTTCTTTAGAGAAACAAATTTCTTGAAGAACGAGTTTCAGAGTATCAGCGTTT 1002
Db 333 IleSerLeuAlaGlyIleThrAsnPhePheGluLysArgValGlyAspTyrGlnLysAla 352
QY 1003 GCAGTTATGGCAGAAAC-----ACAGAT 1026
Db 353 GlyValMetAlaSerThrLysLysAspAlaSerGlnAspThrAlaLysAspAlaAsnAsp 372
QY 1027 AACGCTTCACCTTCGATCCAGATTTT 1053
Db 373 GlyGlyLeuCysPheAspGluAspPhe 381

RESULT 15

US-10-032-585-7537
; Sequence 7537, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7537
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7537

Alignment Scores:
Pred. No.: 2,05e-119 Length: 413
Score: 1192.00 Matches: 225
Percent Similarity: 80.17% Conservative: 58
Best Local Similarity: 63.74% Mismatches: 64
Query Match: 64.57% Indels: 6
DB: 14 Gaps: 3

US-10-698-228-2 (1-1053) x US-10-032-585-7537 (1-413)

QY 13 GAAAGCGGAGCGCGCGGCTGGATCAGATGAGATCATCTTCAGACACCAACGAA 72
Db 61 GluLysAspAspSerLeuLysLysHisGlnAspPheLeuAlaLysHisLysValHisArg 80
QY 73 AGTGAATAAAG-----TCAATGAAGACCTCTTAAGAAAGAGTTCTCGCCGCTTT 126
Db 81 HisLysLeuLysGlnLeuGluAlaGluProLeuValGluAsnLysArgArgTyr 100
QY 127 GTCATCTTTCCCAATCCAGTACCTGATATTTGAAATCGTATAAACAGGCACAGGCTTCC 186
Db 101 ValMetPheProIleArgTyrHisGluIleTrpAsnPheTyrLysLysAlaGluAlaSer 120

QY 187 TTCTGCACGACAGAGGCTCGACTTATCAAAGGATCTCCCTCACTGG---AACAAAGCTT 243
Db 121 PheTrpThrAlaGluGluIleAspLeuSerLysAspLeuAspTrpAsnAsnLysLeu 140
QY 244 AAAGCAGATGAGAAGTACTCTCTCATCTTACATCTTACGCTTTTGGCAGCCAGTATGGA 303
Db 141 AsnGluAsnGluArgTyrPheIleSerArgValLeuAlaPhePheAlaAlaSerAspGly 160
QY 304 ATTGTAATGAATTTGGTGGAGCGCTTTAGTCAGCAGGTGACAGGTTCCAGAGGCTCGC 363
Db 161 IleValGlyGluAsnLeuIleGluAsnPheSerThrGluValGlnLeuProGluAlaLys 180
QY 364 TGTTTCTATGCTTCAAATTCATCATGAGAAATGTTCTCACTCAGAGATGTACAGTTTGTCTG 423
Db 181 SerPheTyrGlyPheGlnIleMetMetGluAsnIleHisSerGluThrTyrSerLeuLeu 200
QY 424 ATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGAAACC 483
Db 201 IleGluThrTyrIleLysAspProGlnGluAlaAspTyrLeuPheAsnAlaIleAlaAsn 220
QY 484 ATGCCCTATGTTAAGAAAGCAGATTTGGCCCTTGCATGGATGATAGCAGATAGAAAATCT 543
Db 221 IleProCysIleGlnLysLysAlaAspTrpAlaIleLysTrpIleGlnAspAspGluAla 240
QY 544 ACTTTTGGGAAAGAGTGGTGGCTTTTCTGCTGTAGAGGAGTTTCTTCTCAGGATCT 603
Db 241 LeuTyrGlyGluArgLeuValAlaPheAlaAlaValGluGlyIlePhePheSerGlySer 260
QY 604 TTTGCTGCTATATTTGGCTTAAAGAGAGAGTCTTATGCCAGGACTCATCTTTTCCAAAT 663
Db 261 PheAlaSerIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSerAsn 280
QY 664 GAATCATCAGCAGATGAGGACTTCACTGTGACTTTCCTGCTGCTGATCTTCCAAATAC 723
Db 281 GluLeuIleCysArgAspGluGlyLeuHisThrAspPheAlaCysLeuLeuPheAlaHis 300
QY 724 TTAGTAAATAAGCTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGTCAAAATTT 783
Db 301 LeuGlnAsnArgProSerProGluIleValGluArgIleIleThrGluAlaValAspVal 320
QY 784 GAGCAGAGTGTTTTAAACAGAAAGCCTTGCAGTTGGCTCTCATTTGGAATGAAATGCAATTTTG 843
Db 321 GluLysGluTyrPheThrAspAlaLeuProValSerLeuLeuGlyMetAsnCysAsnLeu 340
QY 844 ATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTCTCAAG 903
Db 341 MetCysGlnTyrIleGluPheValAlaAspArgLeuLeuAlaLeuGlyAsnLysLys 360
QY 904 GTTTTTCAGGCAGAAATCTTTTGTATTTTATGAAACATTTCTTTAGAGGAAAAACA 963
Db 361 ValTyrAsnValThrAsnProPheAspPheMetGluAsnIleSerLeuAlaGlyLysThr 380
QY 964 AATTTCTTTCAGAAACAGGTTTTCAGAGTATCAGCGTTTTCAGGTTATGGCAGAAACCCACA 1023
Db 381 AsnPhePheGluLysArgValSerAspTyrGlnLysAlaGlyValMetAlaLysThrGlu 400
QY 1024 GAT-----AACGCTTCTCAGCTTGGATCGCAGATTTT 1053
Db 401 AsnLysGluAlaAspAlaPheThrPheAspGluAspPhe 413

Search completed: October 30, 2005, 06:55:05
Job time : 77.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2005, 21:42:15 ; Search time 3258 Seconds
(without alignments)
15660.940 Million cell updates/sec

Title: US-10-698-228-12

Perfect score: 1053

Sequence: 1 atggcgccggaagcc.....tcacttgatgcagatttt 1053

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1053	100.0	1053	6	AR454877 Sequence
2	1053	100.0	1053	6	BD064773 Novel pro
3	1053	100.0	1053	6	BD093086 Novel pro
4	1053	100.0	1081	6	AR454870 Sequence
5	1053	100.0	1081	6	BD064766 Novel pro
6	1053	100.0	1081	6	BD093079 Novel pro
7	1051.4	99.8	1053	6	AR454868 Sequence
8	1051.4	99.8	1053	6	BD064764 Novel pro
9	1051.4	99.8	1053	6	BD093077 Novel pro
10	1051.4	99.8	1056	6	CQ714252 Sequence
11	1051.4	99.8	1056	9	AB166671 Homo sapi
12	1051.4	99.8	4955	6	AR454869 Sequence
13	1051.4	99.8	4955	6	BD064765 Novel pro
14	1051.4	99.8	4955	6	BD093078 Novel pro
15	1051.4	99.8	4955	9	AB036063 Homo sapi
16	1048.2	99.5	1601	6	BD156916 Primer fo
17	1048.2	99.5	1601	6	AX877905 Sequence
18	1048.2	99.5	1601	9	AK001965 Homo sapi
19	896.8	85.2	4532	10	BC058103 Mus muscu

20	853.4	81.0	4571	9	BC042468	Homo sapi
21	851.8	80.9	900	9	AB163437	Homo sapi
22	839.2	79.7	171737	10	AC122379	Mus muscu
23	580.6	55.1	2500	6	AR300103	Sequence
24	580.6	55.1	2500	6	AR438503	Sequence
25	580.6	55.1	2500	6	AR562011	Sequence
26	580	55.1	1794	9	S40301	Sequence
27	580	55.1	2500	6	AR380825	Sequence
28	580	55.1	2500	6	AX333242	Sequence
29	580	55.1	2500	6	AX333687	Sequence
30	580	55.1	2500	6	AX334345	Sequence
31	580	55.1	2500	9	HSRR2SS	X59618 H.sapiens R
32	578.4	54.9	1649	9	BC001886	Homo sapi
33	578.4	54.9	1653	9	BC030154	Homo sapi
34	578.4	54.9	3039	9	AK092671	Homo sapi
35	578.2	54.9	1582	5	BC061353	Xenopus t
36	573.6	54.5	217774	2	AC120603	Rattus no
37	573.2	54.4	3203	9	AK123010	Homo sapi
38	573	54.4	1538	5	CR762018	Xenopus t
39	573	54.4	1611	5	CR848103	Xenopus t
40	573	54.4	1647	5	BC080161	Xenopus t
41	570.4	54.2	2111	10	MUSGRNM2A	M14223 Mouse ribon
42	568.8	54.0	1170	6	E14825	E14825 cDNA encodi
43	566.2	53.8	1570	5	BC072071	Xenopus l
44	564.2	53.6	2057	5	BC044355	Danio rer
45	562.6	53.4	1328	5	DRU57965	Danio rerio

ALIGNMENTS

RESULT 1

AR454877

LOCUS

DEFINITION

AR454877

ACCESSION

AR454877.1

VERSION

GI:42688832

KEYWORDS

Unknown.

SOURCE

Unknown.

ORGANISM

Unclassified.

REFERENCE

1 (bases 1 to 1053)

AUTHORS

Nakamura, Y., Arakawa, H. and Tanaka, H.

TITLE

Protein having a ribonucleotide Reductase activity and a DNA

JOURNAL

Patent: US 6682917-A 12 27-JAN-2004;

FEATURES

Location/Qualifiers

1..1053

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Query Match

Best Local Similarity

Matches 1053; Conservative

100.0%; Score 1053; DB 6; Length 1053;

100.0%; Pred. No. 6.4e-255;

Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGAAAGCGCGGCTGGATCAGATGAGATCATCTTCA 60

1 ATGGGCGACCCGAAAGCGCGGCTGGATCAGATGAGATCATCTTCA 60

QY 61 GACACCAACGAAAGTGAATTAAGTCAANTGAAGCCACTCCTAAGAAAGTCTCGC 120

61 GACACCAACGAAAGTGAATTAAGTCAANTGAAGCCACTCCTAAGAAAGTCTCGC 120

QY 121 CGTTTGTTCATCTTCCATCCAGTACCCCTGATATTTGGAAAAATGTTAAACAGGCACAG 180

121 CGTTTGTTCATCTTCCATCCAGTACCCCTGATATTTGGAAAAATGTTAAACAGGCACAG 180

QY 181 GCTTCTCTTGGACAGCAGAGAGTTGACTTATCAAAAGGATCTCCCTCACTGGAACAAG 240

181 GCTTCTCTTGGACAGCAGAGAGTTGACTTATCAAAAGGATCTCCCTCACTGGAACAAG 240

QY 241 CTTAAGACGATGAGAGAGTACTTCTCTCATCTTAGCCTTTTGGAGCCAGCATGAT 300

[illegible]

```
Db 661 AATGAACATCATCAGCAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Qy 721 TACTTAGTAATAAGCCTTCAGAGAAAGGTCAGGAGATCACTTGTGATGCTGCACAA 780
Db 721 TACTTAGTAATAAGCCTTCAGAGAAAGGTCAGGAGATCACTTGTGATGCTGCACAA 780
Qy 781 ATTGACGAGAGTTTTTAACAGAAAGCCTTGCAGATTTGGCTCATTTGGATGAATTCATT 840
Db 781 ATTGACGAGAGTTTTTAACAGAAAGCCTTGCAGATTTGGCTCATTTGGATGAATTCATT 840
Qy 841 TTGATGAACAGTACATGAGTTTGTAGTGCAGATTAATCTTGTGAACTTGGATTTCTCA 900
Db 841 TTGATGAACAGTACATGAGTTTGTAGTGCAGATTAATCTTGTGAACTTGGATTTCTCA 900
Qy 901 AAGGTTTTTCAGGCAGAAAATCCTTTGATTTATGAGAAACATTTCTTTAGAGGAAA 960
Db 901 AAGGTTTTTCAGGCAGAAAATCCTTTGATTTATGAGAAACATTTCTTTAGAGGAAA 960
Qy 961 ACAATTTCTTTGAGAAAGGTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
Db 961 ACAATTTCTTTGAGAAAGGTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
Qy 1021 ACAGATAACGCTCTTCACTTGGATGCGAGATTTT 1053
Db 1021 ACAGATAACGCTCTTCACTTGGATGCGAGATTTT 1053

RESULT 3
BD093086
LOCUS BD093086 1053 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel protein and its DNA.
ACCESSION BD093086
VERSION BD093086.1 GI:22638674
KEYWORDS WO 0100799-A/10.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1053)
Nakamura, Y., Arakawa, H. and Tanaka, H.
Novel protein and its DNA
Patent: WO 0100799-A 10 04-JAN-2001;
TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,
HIROSHI TANAKA
OS Homo sapiens (human)
PN WO 0100799-A/10
PD 04-JAN-2001
PF 27-JUN-2000 WO 2000JP004189
PR 28-JUN-1999 JP 99P 181131, 06-JUL-1999 JP 99P 192391 PR
21-JAN-2000 JP 00P 017770
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC
.C12P21/02, A61K38/44.
PC A61K48/00, C07K16/40, G01N33/53, C12Q1/26
CC
PH Key Location/Qualifiers.
source 1..1053
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
Query Match 100.0%; Score 1053; DB 6; Length 1053;
Best Local Similarity 100.0%; Pred. No. 6.4e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
Qy 1 ATGGCGGACCCGGAAAGCGCGGAGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGCGGACCCGGAAAGCGCGGAGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Qy 61 GACACCAAGAAAGTGAATAAAGTCAATGAAGAGCCACTCCTAAGAAAGAGTTCGCG 120
```

```
Db 61 GACACCAAGAAAGTGAATAAAGTCAATGAAGAGCCACTCCTAAGAAAGAGTTCGCG 120
Qy 121 CGGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAAATGTATATAACAGGCACAG 180
Db 121 CGGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAAATGTATATAACAGGCACAG 180
Qy 181 GCCTTCTTCTGGACAGCAGAAAGAGTTCATCAAAAGGATCTCCCTCACTGGAACAAG 240
Db 181 GCCTTCTTCTGGACAGCAGAAAGAGTTCATCAAAAGGATCTCCCTCACTGGAACAAG 240
Qy 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTACGCTTTTTCAGGCCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTACGCTTTTTCAGGCCAGTGAT 300
Qy 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGAGTTCAGAGGCT 360
Db 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGAGTTCAGAGGCT 360
Qy 361 CGCTGTTTCTATGCTTCAAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTG 420
Db 361 CGCTGTTTCTATGCTTCAAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTCAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTCAA 480
Qy 481 ACCATGCCCTATCTTAAGAAAAGAGATGGCGCTTGGATGGATAGCAGATAGAAA 540
Db 481 ACCATGCCCTATCTTAAGAAAAGAGATGGCGCTTGGATGGATAGCAGATAGAAA 540
Qy 541 TCTACTTTTGGGAAAAGAGTGGCTTGGCTGTAGAGAGTTCCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAAGAGTGGCTTGGCTGTAGAGAGTTCCTTCTCAGGA 600
Qy 601 TCTTTTGTGCTATATTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACATTTTCC 660
Db 601 TCTTTTGTGCTATATTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACATTTTCC 660
Qy 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 720
Db 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 720
Qy 721 TACTTAGTAATAAGCCTTCAGAGAAAGGTCAGGAGATCACTTGTGATGCTGCACAA 780
Db 721 TACTTAGTAATAAGCCTTCAGAGAAAGGTCAGGAGATCACTTGTGATGCTGCACAA 780
Qy 781 ATTGACGAGAGTTTTTAACAGAAAGCCTTGCAGATTTGGCTCATTTGGATGAATTCATT 840
Db 781 ATTGACGAGAGTTTTTAACAGAAAGCCTTGCAGATTTGGCTCATTTGGATGAATTCATT 840
Qy 841 TTGATGAACAGTACATGAGTTTGTAGTGCAGATTAATCTTGTGAACTTGGATTTCTCA 900
Db 841 TTGATGAACAGTACATGAGTTTGTAGTGCAGATTAATCTTGTGAACTTGGATTTCTCA 900
Qy 901 AAGGTTTTTCAGGCAGAAAATCCTTTGATTTATGAGAAACATTTCTTTAGAGGAAA 960
Db 901 AAGGTTTTTCAGGCAGAAAATCCTTTGATTTATGAGAAACATTTCTTTAGAGGAAA 960
Qy 961 ACAATTTCTTTGAGAAAGGTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
Db 961 ACAATTTCTTTGAGAAAGGTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
Qy 1021 ACAGATAACGCTCTTCACTTGGATGCGAGATTTT 1053
Db 1021 ACAGATAACGCTCTTCACTTGGATGCGAGATTTT 1053

RESULT 4
AR454870
LOCUS AR454870 1081 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 4 from patent US 6682917.
ACCESSION AR454870
```



```
VERSION AR454870.1 GI:42688825
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1081)
AUTHORS Nakamura, Y., Arakawa, H. and Tanaka, H.
TITLE Protein having a ribonucleotide Reductase activity and a DNA
thereof
JOURNAL Patent: US 6682917-A 4 27-JAN-2004;
FEATURES
    source      Location/Qualifiers
                1..1081
                /organism="unknown"
                /mol_type="genomic DNA"
ORIGIN
Query Match      100.0%; Score 1053; DB 6; Length 1081;
Best Local Similarity 100.0%; Pred. No. 6.4e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGCGACCCGGAAGCGCGAAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 60
DB 20 ATGGGCGACCCGGAAGCGCGAAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 79
QY 61 GACACCAACGAAAGTGAAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
DB 80 GACACCAACGAAAGTGAAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 139
QY 121 CGTTTGTGATCTTTCCATTCGATCCCTGATATTTGGAATGATATTAACAGGACAG 180
DB 140 CGTTTGTGATCTTTCCATTCAGTACCCTGATATTTGGAATGATATTAACAGGACAG 199
QY 181 GCTTCTCTCTGACACGACGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
DB 200 GCTTCTCTCTGACACGACGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 259
QY 241 CTTAAGCAGATGAGAAGTACTTCTCTCACAATCTTAGCCCTTTTTCGAGCCAGTGAT 300
DB 260 CTTAAGCAGATGAGAAGTACTTCTCTCACAATCTTAGCCCTTTTTCGAGCCAGTGAT 319
QY 301 GGAATGTAAATGAAAATTTGGTGGAGCGCTTTATGTCAGAGGTGAGGTTCCAGAGGCT 360
DB 320 GGAATGTAAATGAAAATTTGGTGGAGCGCTTTATGTCAGAGGTGAGGTTCCAGAGGCT 379
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAAGTGTCTCATCAGAGATGACAGTTTG 420
DB 380 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAAGTGTCTCATCAGAGATGACAGTTTG 439
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 480
DB 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 499
QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAA 540
DB 500 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAA 559
QY 541 TCTACTTTGGGGAAGAGTGTGGCTTTGCTGCTGTAGAGAGGTTTCTTCTCAGGA 600
DB 560 TCTACTTTGGGGAAGAGTGTGGCTTTGCTGCTGTAGAGAGGTTTCTTCTCAGGA 619
QY 601 TCTTTTGTGCTATATTCTGGCTTAAAGAAAGAGAGGTTCTTATGCGAGGACTCACTTTTTC 660
DB 620 TCTTTTGTGCTATATTCTGGCTTAAAGAAAGAGAGGTTCTTATGCGAGGACTCACTTTTTC 679
QY 661 AATGAACCTCATCAGCAGATGAAGGACTTCACTGTGACTTTTGTGCTTGTGTTCCAA 720
DB 680 AATGAACCTCATCAGCAGATGAAGGACTTCACTGTGACTTTTGTGCTTGTGTTCCAA 739
QY 721 TACTTGTAAATTAAGCTTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
DB 740 TACTTGTAAATTAAGCTTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 799
QY 781 ATTGAGCAGGAGTTTTTAAACAGAAAGCCTTTGCCAGTGTGGCTCATTTGGAATGATTCAT 840
DB 800 ATTGAGCAGGAGTTTTTAAACAGAAAGCCTTTGCCAGTGTGGCTCATTTGGAATGATTCAT 859
QY 841 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTGTGGACTTTGGATTCTCA 900
DB 860 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTGTGGACTTTGGATTCTCA 919
QY 901 AAGGTTTTTTCAGGCAGAAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAAAGGAAA 960
DB 920 AAGGTTTTTTCAGGCAGAAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAAAGGAAA 979
QY 961 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTTCAGTTATGCGAGAAACC 1020
DB 980 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTTCAGTTATGCGAGAAACC 1039
QY 1021 ACAGATAAAGCTCTTACCTTGGATGACAGATTTT 1053
DB 1040 ACAGATAAAGCTCTTACCTTGGATGACAGATTTT 1072
RESULT 5
BD064766 LOCUS 1081 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel protein and DNA thereof.
ACCESSION BD064766
VERSION 1 GI:22610369
KEYWORDS JP 2001269184-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1081)
AUTHORS Nakamura, Y., Arakawa, H. and Tanaka, H.
TITLE Novel protein and DNA thereof
JOURNAL Patent: JP 2001269184-A 3 02-OCT-2001;
YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD
COMMENT OS Homo sapiens (human)
PN JP 2001269184-A/3
PD 02-OCT-2001
PF 27-JUN-2000 JP 2000192401
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N15/09, A61K31/711, A61K38/00, A61K38/43, A61K45/00, A61K48/00, PC
A61P35/00,
PC A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC
C12N5/10, C12N9/02,
PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53// (C12N15/09,
C12R1:91),
PC (C12N15/09, C12R1:19), (C12N9/02, C12R1:91), (C12P21/02, C12R1:91),
PC (C12P21/08, C12R1:91), (C12N15/00, A61K37/02, A61K37/48, C12N5/00,
CC (C12N15/00, C12R1:91), (C12N15/00, C12R1:19)
CC Novel protein and DNA thereof
FH Key Location/Qualifiers
FT source 1..1081
FT /organism="Homo sapiens (human)".
FEATURES
    source      Location/Qualifiers
                1..1081
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
ORIGIN
Query Match      100.0%; Score 1053; DB 6; Length 1081;
Best Local Similarity 100.0%; Pred. No. 6.4e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGCGACCCGGAAGCGCGAAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 60
DB 20 ATGGGCGACCCGGAAGCGCGAAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 79
QY 61 GACACCAACGAAAGTGAAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
DB 80 GACACCAACGAAAGTGAAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 139
QY 121 CGTTTGTGATCTTTCCATTCGATCCCTGATATTTGGAATGATATTAACAGGACAG 180
DB 140 CGTTTGTGATCTTTCCATTCAGTACCCTGATATTTGGAATGATATTAACAGGACAG 199
QY 181 GCTTCTCTCTGACACGACGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
DB 200 GCTTCTCTCTGACACGACGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 259
QY 241 CTTAAGCAGATGAGAAGTACTTCTCTCACAATCTTAGCCCTTTTTCGAGCCAGTGAT 300
DB 260 CTTAAGCAGATGAGAAGTACTTCTCTCACAATCTTAGCCCTTTTTCGAGCCAGTGAT 319
QY 301 GGAATGTAAATGAAAATTTGGTGGAGCGCTTTATGTCAGAGGTGAGGTTCCAGAGGCT 360
DB 320 GGAATGTAAATGAAAATTTGGTGGAGCGCTTTATGTCAGAGGTGAGGTTCCAGAGGCT 379
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAAGTGTCTCATCAGAGATGACAGTTTG 420
DB 380 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAAGTGTCTCATCAGAGATGACAGTTTG 439
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 480
DB 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 499
QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAA 540
DB 500 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAA 559
QY 541 TCTACTTTGGGGAAGAGTGTGGCTTTGCTGCTGTAGAGAGGTTTCTTCTCAGGA 600
DB 560 TCTACTTTGGGGAAGAGTGTGGCTTTGCTGCTGTAGAGAGGTTTCTTCTCAGGA 619
QY 601 TCTTTTGTGCTATATTCTGGCTTAAAGAAAGAGAGGTTCTTATGCGAGGACTCACTTTTTC 660
DB 620 TCTTTTGTGCTATATTCTGGCTTAAAGAAAGAGAGGTTCTTATGCGAGGACTCACTTTTTC 679
QY 661 AATGAACCTCATCAGCAGATGAAGGACTTCACTGTGACTTTTGTGCTTGTGTTCCAA 720
DB 680 AATGAACCTCATCAGCAGATGAAGGACTTCACTGTGACTTTTGTGCTTGTGTTCCAA 739
QY 721 TACTTGTAAATTAAGCTTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
DB 740 TACTTGTAAATTAAGCTTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 799
QY 781 ATTGAGCAGGAGTTTTTAAACAGAAAGCCTTTGCCAGTGTGGCTCATTTGGAATGATTCAT 840
```


[illegible]

RESULT 7
 AR454868
 LOCUS
 DEFINITION
 Sequence 2 from patent US 6682917.
 AR454868
 ACCESSION
 AR454868.1 GI:42688823
 VERSION
 KEYWORDS
 SOURCE
 Unknown.
 ORGANISM
 Unknown.
 UNCLASSIFIED.
 REFERENCE
 1 (bases 1 to 1053)
 AUTHORS
 Nakamura, Y., Arakawa, H. and Tanaka, H.
 TITLE
 Protein having a ribonucleotide reductase activity and a DNA
 thereof
 JOURNAL
 Patent: US 6682917-A 2 27-JAN-2004;
 FEATURES
 Location/Qualifiers
 1..1053
 /organism="unknown"
 /mol_type="genomic DNA"
 source

Query Match	99.8%	Score 1051.4	DB 6	Length 1053
Best Local Similarity	99.9%	Pred. No. 1.6e-254		
Matches 1052: Conservative	0	Mismatches 1	Indels 0	Gaps 0

Qy	1	ATGGGGACCCGGAAAGCCCGGAAAGCCCGGGCTGGATCAGGATGAGAGATCATCTTCA	60
Db	1	ATGGGGACCCGGAAAGCCCGGAAAGCCCGGGCTGGATCAGGATGAGAGATCATCTTCA	60
Qy	61	GACACCAACGAAGTGAATAAAGTCAAAATGAAGAGCCACCTCTTAAAGAAAGAGTTCTCGC	120
Db	61	GACACCAACGAAGTGAATAAAGTCAAAATGAAGAGCCACCTCTTAAAGAAAGAGTTCTCGC	120
Qy	121	CGGTTTGTCATCTTTTCCAAATCCAGTACCCCTGATATTTGGAAAAATGTTATAAACAGGSCACAG	180
Db	121	CGGTTTGTCATCTTTTCCAAATCCAGTACCCCTGATATTTGGAAAAATGTTATAAACAGGSCACAG	180
Qy	181	GCTTCTCTTGGACAGCAGAAAGAGTTGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG	240
Db	181	GCTTCTCTTGGACAGCAGAAAGAGTTGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG	240

[illegible]

RESULT	8
LOCUS	BD064764
DEFINITION	Novel protein and DNA thereof.
ACCESSION	BD064764
VERSION	BD064764.1 GI:22610367
KEYWORDS	JP 2001269184-A/1.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 1053)
AUTHORS	Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE	Novel protein and DNA thereof
JOURNAL	Patent: JP 2001269184-A 1 02-OCT-2001; YUSUKE NAKAMURA,TAKEDA CHEMICAL INDUSTRIES LTD

```
COMMENT OS Homo sapiens (human)
PN JP 2001269184-A/1
PD 02-OCT-2001
PF 27-JUN-2000 JP 2000192401
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N15/09, A61K31/11, A61K38/00, A61K38/43, A61K45/00, A61K48/00, PC
A61P35/00,
PC A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC
, C12N5/10, C12N5/02,
PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53// (C12N15/09,
C12R1/91),
PC (C12N15/09, C12R1/19), (C12N9/02, C12R1/91), (C12P21/02, C12R1/91),
PC (C12P21/08, C12R1/91), (C12N15/00, A61K37/02, A61K37/48, C12N5/00,
PC (C12N15/00, C12R1/91), (C12N15/00, C12R1/19)
CC Novel protein and DNA thereof
FH Key Location/Qualifiers
FT source 1..1053
FT /organism='Homo sapiens (human)'.
FEATURES
source
1..1053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 1053;
Best Local Similarity 99.9%; Pred. No. 1.6e-254;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGCGACCCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 1 ATGGCGACCCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACCAAGCAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 61 GACACCAAGCAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 121 CGGTTTGTCATCTTCCATCCAGTACCTGATATTTGGAATGATTAACAGGCACAG 180
DB 121 CGGTTTGTCATCTTCCATCCAGTACCTGATATTTGGAATGATTAACAGGCACAG 180
QY 181 GCTTCCTCTTGACACGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
DB 181 GCTTCCTCTTGACACGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
QY 241 CTTAAGCAGATGAGAAGTACTTCACTCTCAATCTTAGCCCTTTTTCAGCCAGTGAT 300
DB 241 CTTAAGCAGATGAGAAGTACTTCACTCTCAATCTTAGCCCTTTTTCAGCCAGTGAT 300
QY 301 GGAATGTAATGAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
DB 301 GGAATGTAATGAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
QY 361 CGCTGTTCTATGGCTTTTCAAAATCTCATCGAGAATGTTTCACTCAGAGATGACAGTTG 420
DB 361 CGCTGTTCTATGGCTTTTCAAAATCTCATCGAGAATGTTTCACTCAGAGATGACAGTTG 420
QY 421 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 480
DB 421 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCCCTTCGATGGATAGCAGATAGAAA 540
DB 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCCCTTCGATGGATAGCAGATAGAAA 540
QY 541 TCTACTTTGGGGAAGAGTGTGGCTTTTGTGCTGTAGAGGAGTTTCTTCTCAGGA 600
DB 541 TCTACTTTGGGGAAGAGTGTGGCTTTTGTGCTGTAGAGGAGTTTCTTCTCAGGA 600
QY 601 TCTTTTGTGCTATATTCTGGCTAAAGAAAGAGAGTCTTATGCGCAGACTCACTTTTCC 660
DB 601 TCTTTTGTGCTATATTCTGGCTAAAGAAAGAGAGTCTTATGCGCAGACTCACTTTTCC 660
```

```
QY 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGTGACTTTGCTTGCTGTGATTTCCAA 720
DB 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGTGACTTTGCTTGCTGTGATTTCCAA 720
QY 721 TACTTAGTAATAAGCCCTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTGCAAA 780
DB 721 TACTTAGTAATAAGCCCTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTGCAAA 780
QY 781 ATTGAGCAGAGGTTTTTAAACAGAAAGCCCTTGGCCAGTTGGCCCTCATTTGGAATGAATTCGATT 840
DB 781 ATTGAGCAGAGGTTTTTAAACAGAAAGCCCTTGGCCAGTTGGCCCTCATTTGGAATGAATTCGATT 840
QY 841 TTGATGAACACAGTACATTTAGTTTGTAGCTGACAGATTTACTTGTGGAATTTGGATTTCTCA 900
DB 841 TTGATGAACACAGTACATTTAGTTTGTAGCTGACAGATTTACTTGTGGAATTTGGATTTCTCA 900
QY 901 AAGTTTTTTCAGGAGAAAAATCCTTTTGAATTTTATGGAACATTTCTTTAGAGAAAAA 960
DB 901 AAGTTTTTTCAGGAGAAAAATCCTTTTGAATTTTATGGAACATTTCTTTAGAGAAAAA 960
QY 961 ACAATTTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTTGGCAGTTTGGCAGAAACC 1020
DB 961 ACAATTTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTTGGCAGTTTGGCAGAAACC 1020
QY 1021 ACAGATAACGCTTCTTCACTTTGGATGAGATTTT 1053
DB 1021 ACAGATAACGCTTCTTCACTTTGGATGAGATTTT 1053
RESULT 9
BD093077
LOCUS BD093077 1053 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel protein and its DNA.
ACCESSION BD093077
VERSION BD093077.1 GI:22638665
KEYWORDS WO 0100799-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1053)
AUTHORS Nakamura, Y., Arakawa, H. and Tanaka, H.
TITLE Novel protein and its DNA
JOURNAL Patent: WO 0100799-A 1 04-JAN-2001;
TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,
HIROSHI TANAKA
COMMENT OS Homo sapiens (human)
PN WO 0100799-A/1
PD 04-JAN-2001
PF 27-JUN-2000 WO 2000P004189
PR 28-JUN-1999 JP 99P 181131, 06-JUL-1999 JP 99P 192391 PR
21-JAN-2000 JP 00P 017770
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC
, C12P21/02, A61K38/44,
PC A61K48/00, C07K16/40, G01N33/53, C12Q1/26
CC
FH Key Location/Qualifiers.
FEATURES
source
1..1053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 1053;
Best Local Similarity 99.9%; Pred. No. 1.6e-254;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGCGACCCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 1 ATGGCGACCCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
```

```
QY 61 GACACCAAGAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAAGAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 121 CGGTTTGTCTATCTTTCCATCCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG 180
Db 121 CGGTTTGTCTATCTTTCCATCCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG 180
QY 181 GCTTCTCTCTGGACAGACAGAGGTGACTTATCAAAAGATCTCCCTCACTGGAACAAG 240
Db 181 GCTTCTCTCTGGACAGACAGAGGTGACTTATCAAAAGATCTCCCTCACTGGAACAAG 240
QY 241 CTTAAAGCAGATGAGAAGTACTCATCTCTCAATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAAGTACTCATCTCTCAATCTTAGCCCTTTTTCAGCCAGTGAT 300
QY 301 GGAATTTGAATGAATAATTTGGTGGAGCCCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Db 301 GGAATTTGAATGAATAATTTGGTGGAGCCCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
QY 481 ACATGCCCCTATGTTAAGAAAAAAGCAGATTTGGGCCCTTCGATGGATGAGATAGAAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCCCTTCGATGGATGAGATAGAAAA 540
QY 541 TCTACTTTTGGGAAAGAGTGTGGCCTTTGCTGTGTAGAGGAGTCTTCTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGTGGCCTTTGCTGTGTAGAGGAGTCTTCTCTCAGGA 600
QY 601 TCTTTTGTCTGTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
Db 601 TCTTTTGTCTGTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
QY 661 ATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
QY 721 TACTTAGTAAATAGCTTTCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 721 TACTTAGTAAATAGCTTTCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
QY 781 ATTGAGCAGGAGTTTAAACAGAGGCTTGGCAGTTGGCTCATTTGGAATGGAATTCATT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGGCTTGGCAGTTGGCTCATTTGGAATGGAATTCATT 840
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGTGAAGATTAATTGTTGGAATCTGATTTCTCA 900
Db 841 TTGATGAAACAGTACATTCAGTTTGTAGTGAAGATTAATTGTTGGAATCTGATTTCTCA 900
QY 901 AAGGTTTTTCAGCAGAAAAATCCTTTTGAATTTATGGAATAACATTCTTTAGAGAAAA 960
Db 901 AAGGTTTTTCAGCAGAAAAATCCTTTTGAATTTATGGAATAACATTCTTTAGAGAAAA 960
QY 961 ACAAAATTTCTTTGAGAAAACAGTTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAAACAGTTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
QY 1021 ACAGATAACGCTTTCACCTTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGCTTTCACCTTTGGATGAGATTTT 1053
```

RESULT 10
CQ714252
LOCUS
DEFINITION Sequence 186 from Patent WO02068579.
QY14252 1056 bp DNA linear PAT 03-FEB-2004

```
ACCESSION CQ714252
VERSION CQ714252.1
KEYWORDS GI:42275109
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
JOURNAL Patent: WO 02068579-A 186 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
    source
        1..1056
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 1056;
Best Local Similarity 99.9%; Pred. No. 1.6e-254;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGGCGACCCCGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGGCGACCCCGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTCTTAAGAAAGAGTTCTCGC 120
QY 121 CGGTTTGTCTATCTTTCCATCCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG 180
Db 121 CGGTTTGTCTATCTTTCCATCCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG 180
QY 181 GCTTCTCTCTGGACAGCAGAGAGGTGACTTATCAAAAGATCTCCCTCACTGGAACAAG 240
Db 181 GCTTCTCTCTGGACAGCAGAGAGGTGACTTATCAAAAGATCTCCCTCACTGGAACAAG 240
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCAATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCAATCTTAGCCCTTTTTCAGCCAGTGAT 300
QY 301 GGAATTTGAATGAATAATTTGGTGGAGCCCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Db 301 GGAATTTGAATGAATAATTTGGTGGAGCCCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCCCTTCGATGGATGAGATAGAAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCCCTTCGATGGATGAGATAGAAAA 540
QY 541 TCTACTTTTGGGAAAGAGTGTGGCCTTTGCTGTGTAGAGGAGTCTTCTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGTGGCCTTTGCTGTGTAGAGGAGTCTTCTCTCAGGA 600
QY 601 TCTTTTGTCTGTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
Db 601 TCTTTTGTCTGTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
```

QY 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAA 780
 Db 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAA 780
 QY 781 ATTGACGAGGAGTTTAAACAGAGAGCTTGCAGTTGGCTCATTTGGAATGAATTGCATT 840
 Db 781 ATTGACGAGGAGTTTAAACAGAGAGCTTGCAGTTGGCTCATTTGGAATGAATTGCATT 840
 QY 841 TTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTAATTCTGTGAACTTTGGATTCTCA 900
 Db 841 TTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTAATTCTGTGAACTTTGGATTCTCA 900
 QY 901 AAGGTTTTTTCAGCAGAGAAATCCTTTTGTATTTATGGAACACATTTCTTTAGAGGAAAA 960
 Db 901 AAGGTTTTTTCAGCAGAGAAATCCTTTTGTATTTATGGAACACATTTCTTTAGAGGAAAA 960
 QY 961 ACAAAATTTCTTGGAGAAACAGTTTCAGAGTATCAGGTTTTCAGAGTATGCGAGAAACC 1020
 Db 961 ACAAAATTTCTTGGAGAAACAGTTTCAGAGTATCAGGTTTTCAGAGTATGCGAGAAACC 1020
 QY 1021 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1053
 Db 1021 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1053

RESULT 11
 AB166671 1056 bp mRNA linear PRI 09-MAR-2004
 LOCUS Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase
 DEFINITION small subunit 2, complete cds.
 ACCESSION AB166671
 VERSION 1
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
 Ugai, H. and Yokoyama, K.K.
 AUTHORS Ugai, H. and Yokoyama, K.K.
 TITLE Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase
 JOURNAL Published Only in Database (2004)
 REFERENCE 2 (bases 1 to 1056)
 AUTHORS Ugai, H. and Yokoyama, K.K.
 TITLE Direct Submission
 JOURNAL Submitted (05-MAR-2004) Hideyo Ugai, RIKEN BioResource Center, Gene
 Engineering Division; 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074,
 Japan [E-mail: ugai@rtc.riken.go.jp, Tel: 81-29-836-3612,
 Fax: 81-29-836-9120]

FEATURES
 source location/Qualifiers
 1..1056
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="p53R2.C19"
 /cell_line="Saos-2"
 /cell_type="Osteogenic sarcoma"
 1..1056
 /gene="p53R2"
 1..1056
 /gene="p53R2"
 /codon_start=1
 /product="p53-inducible ribonucleotide reductase small
 subunit 2"
 /protein_id="BAD12267.1"
 /db_xref="GI:45259569"
 /translation="MGDPERPEAAGLDQDERSSTNSEIKNSEPELRLKSRPFVI
 FPIQPIWKMTQKQASFWTAEEVDLSKDLPHWNKLKADKFIYFISHILAFPAASDGI
 VNEMLVFSQVQPEARCFYQFQILNENHSEMYSLIDYIIRDPKKREFLNAIE
 TMYVYKKADWALRIADRKSTFGERVVAFAAAGVFFSGSFAAIFWLKKRGLMPLGT
 FSNELISDEGLHCDIFACLMFOYLVNKSEERVREIIVDAVKIEOEFLEALPVGLIG
 MNCILMKOYIEFVADRLVLELGFSGVQENPDPFENISLSEGTNFFEKRVSEVQRF
 AVMAETTNVFLDADF"

ORIGIN

Query Match 99.8%; Score 1051.4; DB 9; Length 1056;
 Best Local Similarity 99.9%; Pred. No. 1.6e-254;
 Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGCAGCCCGGAAAGCGCGGAGCGCGGCTGGATCAGAGATCATCTTCA 60
 Db 1 ATGGCGCAGCCCGGAAAGCGCGGAGCGCGGCTGGATCAGAGATCATCTTCA 60
 QY 61 GACACCAACGAAAGTGAATAAATCAAAATGAAGAGCCACTCTTAAGAAAGATTCTCG 120
 Db 61 GACACCAACGAAAGTGAATAAATCAAAATGAAGAGCCACTCTTAAGAAAGATTCTCG 120
 QY 121 CGGTTGTCTATCTTTCCAAATCCAGTACCTGATATTTGGAAATCTATAAACAGGCACAG 180
 Db 121 CGGTTGTCTATCTTTCCAAATCCAGTACCTGATATTTGGAAATCTATAAACAGGCACAG 180
 QY 181 GCTTCTCTTCTGGACAGAGAGGTTGACTTATCAAAAGGATCTCCCTCAGTGGAAACAG 240
 Db 181 GCTTCTCTTCTGGACAGAGAGGTTGACTTATCAAAAGGATCTCCCTCAGTGGAAACAG 240
 QY 241 CTTAAAGCAGATGAGAAATCTTCAATCTCTCAATCTTTAGCCCTTTTTCAGCCAGTGAT 300
 Db 241 CTTAAAGCAGATGAGAAATCTTCAATCTCTCAATCTTTAGCCCTTTTTCAGCCAGTGAT 300
 QY 301 GGAATTTGTAATGAAATTTGGTGGAGGCTTTTAGTACAGGAGTCCAGGCTCCAGAGCT 360
 Db 301 GGAATTTGTAATGAAATTTGGTGGAGGCTTTTAGTACAGGAGTCCAGGCTCCAGAGCT 360
 QY 361 CGCTGTTTCTATGCTTTCAAAATTTCTATCAGAAATGTTCACTCAGAGATGACAGTTTG 420
 Db 361 CGCTGTTTCTATGCTTTCAAAATTTCTATCAGAAATGTTCACTCAGAGATGACAGTTTG 420
 QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTAATCAATTTGAA 480
 Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTAATCAATTTGAA 480
 QY 481 ACCATGCCCTATGTTAAAGAAAAAGAGATTGGGCTTGGCGATGATGATAGAGAAAA 540
 Db 481 ACCATGCCCTATGTTAAAGAAAAAGAGATTGGGCTTGGCGATGATGATAGAGAAAA 540
 QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTCTCTCTGTAGAGAGGATTTCTTCTCAGGA 600
 Db 541 TCTACTTTTGGGGAAGAGTGGTGGCTTCTCTCTGTAGAGAGGATTTCTTCTCAGGA 600
 QY 601 TCTTTTGTCTCTATATTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCATTCTTTTCC 660
 Db 601 TCTTTTGTCTCTATATTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCATTCTTTTCC 660
 QY 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGTGCTGCTGATGTCCAA 720
 Db 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGTGCTGCTGATGTCCAA 720
 QY 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAA 780
 Db 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAA 780
 QY 781 ATTGAGCAGGAGTTTAAACAGAGAGCTTGCAGTTGGCCCTCATTTGGAATGAATTGCATT 840
 Db 781 ATTGAGCAGGAGTTTAAACAGAGAGCTTGCAGTTGGCCCTCATTTGGAATGAATTGCATT 840
 QY 841 TTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTAATTCTGTGAACTTTGGATTCTCA 900
 Db 841 TTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTAATTCTGTGAACTTTGGATTCTCA 900
 QY 901 AAGGTTTTTTCAGCAGAGAAATCCTTTTGTATTTATGGAACACATTTCTTTAGAGAAAA 960
 Db 901 AAGGTTTTTTCAGCAGAGAAATCCTTTTGTATTTATGGAACACATTTCTTTAGAGAAAA 960
 QY 961 ACAAAATTTCTTGGAGAAACAGTTTCAGAGTATCAGGTTTTCAGAGTATGCGAGAAACC 1020
 Db 961 ACAAAATTTCTTGGAGAAACAGTTTCAGAGTATCAGGTTTTCAGAGTATGCGAGAAACC 1020

Best Local Similarity 99.9%; Pred. No. 1.5e-254; Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;									
QY	1	ATGGCGACCCGGAAGCGCGGAGCGGCTGGATCAGGATCAGAGATCATCTTCA	60						
Db	245	ATGGCGACCCGGAAGCGGAGCGGCTGGATCAGGATCAGAGATCATCTTCA	304						
QY	61	GACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC	120						
Db	305	GACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC	364						
QY	121	CGTTTGTCTATCTTCCATTCAGTACCTGATATTTGGAAATGATATAACAGGCACAG	180						
Db	365	CGTTTGTCTATCTTCCATTCAGTACCTGATATTTGGAAATGATATAACAGGCACAG	424						
QY	181	GCTTCTCTCTGACAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGACAG	240						
Db	425	GCTTCTCTCTGACAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGACAG	484						
QY	241	CTTAAAGCAGATCAGAGAGTACTTCTCATCATCTTTAGCCTTTTTCAGCCAGTGTAT	300						
Db	485	CTTAAAGCAGATCAGAGAGTACTTCTCATCATCTTTAGCCTTTTTCAGCCAGTGTAT	544						
QY	301	GGAAATGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGCTTCAGAGGCT	360						
Db	545	GGAAATGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGCTTCAGAGGCT	604						
QY	361	CGCTGTTTCTATGCTTTTCAAAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTTG	420						
Db	605	CGCTGTTTCTATGCTTTTCAAAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTTG	664						
QY	421	CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA	480						
Db	665	CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA	724						
QY	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAAA	540						
Db	725	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAAA	784						
QY	541	TCTACTTTTGGGAAAGAGTGGGCTTGTCTGTGTAGAGAGTGTTCCTCAGGA	600						
Db	785	TCTACTTTTGGGAAAGAGTGGGCTTGTCTGTGTAGAGAGTGTTCCTCAGGA	844						
QY	601	TCTTTTGTCTATATCTTGGCTTAAGAAAGAGAGGCTTATGTCAGGACTCATCTTTTCC	660						
Db	845	TCTTTTGTCTATATCTTGGCTTAAGAAAGAGAGGCTTATGTCAGGACTCATCTTTTCC	904						
QY	661	AATGAATCTCATCAGCAGAGATGAAGACTTCTACTGTGACTTTTCTGCTTGGCTGATTTCCAA	720						
Db	905	AATGAATCTCATCAGCAGAGATGAAGACTTCTACTGTGACTTTTCTGCTTGGCTGATTTCCAA	964						
QY	721	TACTTAGTAAATGAAGCTTTCAGAAAGAGGTCAGGAGATCATTTGTGATGCTGTCAAA	780						
Db	965	TACTTAGTAAATGAAGCTTTCAGAAAGAGGTCAGGAGATCATTTGTGATGCTGTCAAA	1024						
QY	781	ATTGAGCAGAGTTTATCAGAGCCTTGCAGTGGCTCTATGGAATGGAATTCGAT	840						
Db	1025	ATTGAGCAGAGTTTATCAGAGCCTTGCAGTGGCTCTATGGAATGGAATTCGAT	1084						
QY	841	TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTGTGAACTTGGATTTCTCA	900						
Db	1085	TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTGTGAACTTGGATTTCTCA	1144						
QY	901	AAGGTTTTTTCAGGCAGAAATCCTTTTGTATTTTATGGAACACATTTCTTTAGAGGAAA	960						
Db	1145	AAGGTTTTTTCAGGCAGAAATCCTTTTGTATTTTATGGAACACATTTCTTTAGAGGAAA	1204						
QY	961	ACAAATTTCTTTGAGAAAACAGTTTTCAGAGTATCAGCGTTTTCAGGATGTACAGTTTG	1020						
Db	1205	ACAAATTTCTTTGAGAAAACAGTTTTCAGAGTATCAGCGTTTTCAGGATGTACAGTTTG	1264						
QY	1021	ACAGATAAGCTTCTACCTTTGGATGAGATTTT 1053							

Db	1265	ACAGATAAGCTTCTACCTTGGATGCAGATTTT 1297							
RESULT 14									
LOCUS	BD093078	Novel protein and its DNA.	4955 bp	DNA	linear	PAT 27-AUG-2002			
DEFINITION	BD093078	Novel protein and its DNA.							
ACCESSION	BD093078	Novel protein and its DNA							
VERSION	BD093078.1	GI:22638666							
KEYWORDS	WO 0100799-A/2;								
SOURCE	Homo sapiens (human)								
ORGANISM	Homo sapiens								
REFERENCE	1 (bases 1 to 4955)								
AUTHORS	Nakamura,Y., Arakawa,H. and Tanaka,H.								
TITLE	Novel protein and its DNA								
JOURNAL	Patent: WO 0100799-A 2 04-JAN-2001; TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA								
COMMENT	OS Homo sapiens (human) PN WO 0100799-A/2 PD 04-JAN-2001 PF 27-JUN-2000 WO 2000JP004189 PR 28-JUN-1999 JP 99P 181131.06-JUL-1999 JP 99P 192391 PR 21-JAN-2000 JP 00P 017770 PI YUSUKE NAKAMURA,HIROFUMI ARAKAWA,HIROSHI TANAKA PC C12N9/04,C12N15/53,C12N1/15,C12N1/19,C12N1/21,C12N5/10 PC ,C12P21/02,A61K38/44, PC A61K48/00,C07K16/40,G01N33/53,C12Q1/26 CC								
FEATURES	Key Location/Qualifiers								
source	1..4955								
	/organism="Homo sapiens"								
	/mol_type="genomic DNA"								
	/db_xref="taxon:9606"								
ORIGIN									
Query Match	99.8%; Score 1051.4; DB 6; Length 4955;								
Best Local Similarity	99.9%; Pred. No. 1.5e-254;								
Matches 1052;	Conservative 0; Mismatches 1; Indels 0; Gaps 0;								
QY	1	ATGGCGACCCGGAAGCGCGGAGCGGCTGGATCAGGATCAGAGATCATCTTCA	60						
Db	245	ATGGCGACCCGGAAGCGCGGAGCGGCTGGATCAGGATCAGAGATCATCTTCA	304						
QY	61	GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC	120						
Db	305	GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC	364						
QY	121	CGTTTGTCTATCTTTCCAAATCCAGTACCTGATATTTGGAAATGATATAACAGGCACAG	180						
Db	365	CGTTTGTCTATCTTTCCAAATCCAGTACCTGATATTTGGAAATGATATAACAGGCACAG	424						
QY	181	GCTTCTCTCTGACAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAG	240						
Db	425	GCTTCTCTCTGACAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAG	484						
QY	241	CTTAAAGCAGATCAGAGAGTACTTCTCATCATCTTTAGCCTTTTTCAGCCAGTGTAT	300						
Db	485	CTTAAAGCAGATCAGAGAGTACTTCTCATCATCTTTAGCCTTTTTCAGCCAGTGTAT	544						
QY	301	GGAAATGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGCTTCAGAGGCT	360						
Db	545	GGAAATGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGCTTCAGAGGCT	604						
QY	361	CGCTGTTTCTATGCTTTTCAAAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTTG	420						
Db	605	CGCTGTTTCTATGCTTTTCAAAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTTG	664						
QY	421	CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA	480						


```
Db      |||||||
905 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
Qy      |||||||
721 TACTTAGTAAATAAGCCTTCAGAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db      |||||||
965 TACTTAGTAAATAAGCCTTCAGAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 1024
Qy      |||||||
781 ATTGAGCAGGAGTTTAAACAGAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTGCATT 840
Db      |||||||
1025 ATTGAGCAGGAGTTTAAACAGAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTGCATT 1084
Qy      |||||||
841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGTGAACTTGGATTCTCA 900
Db      |||||||
1085 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGTGAACTTGGATTCTCA 1144
Qy      |||||||
901 AAGGTTTTTCAGGCAGAAAATCCTTTTGATTTTATGGAACAATTTCTTTAGAGGAAAA 960
Db      |||||||
1145 AAGGTTTTTCAGGCAGAAAATCCTTTTGATTTTATGGAACAATTTCTTTAGAGGAAAA 1204
Qy      |||||||
961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
Db      |||||||
1205 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1264
Qy      |||||||
1021 ACAGATAACGTTCTTACCTTTGGATGCAGATTTT 1053
Db      |||||||
1265 ACAGATAACGTTCTTACCTTTGGATGCAGATTTT 1297
```

Search completed: October 30, 2005, 01:51:31
Job time : 3261 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2005, 21:38:30 ; Search time 440.5 Seconds
(without alignments)
14150.937 Million cell updates/sec

Title: US-10-698-228-12

Perfect score: 1053

Sequence: 1 atggcgccaccggaaggcc.....tcacctggatgcagatttt 1053

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1053	100.0	1053	4 AAF32447	Aaf32447 Human rib
2	1053	100.0	1081	4 AAF32440	Aaf32440 Human rib
3	1051.4	99.8	1053	4 AAF32438	Aaf32438 Human rib
4	1051.4	99.8	2596	3 AAI12411	Aai12411 cDNA enco
5	1051.4	99.8	4955	4 AAF32439	Aaf32439 Human rib
6	1051.4	99.8	4955	13 ADR24210	Adr24210 Breast ca
7	1048.2	99.5	1601	4 AAH14924	Aah14924 Human cDN
8	650.6	61.8	706	13 ADQ57092	Adq57092 Novel can
9	580	55.1	2482	4 AAS444917	Aas44917 Human con
10	580	55.1	2500	6 ABL65414	Ab165414 Lung canc
11	580	55.1	2500	6 ABL66517	Ab166517 Lung canc
12	580	55.1	2500	6 ABL65859	Ab165859 Lung canc
13	580	55.1	2500	8 ABL10335	Ab10335 DNA enco
14	580	55.1	2500	11 ADI32044	Adi32044 Human cDN
15	580	55.1	2500	12 ADN04443	Adn04443 Antipsori
16	580	55.1	2500	12 ADQ09273	Adq09273 Human RRM
17	580	55.1	2500	13 ACN37637	Acn37637 Tumour-as
18	580	55.1	3393	12 ADN03788	Adn03788 Antipsori
19	580	55.1	3393	12 ADO19225	Ado19225 Human PRO
20	578.4	54.9	1989	3 AAC78111	Aac78111 Human can

ALIGNMENTS

RESULT 1

AAF32447

ID AAF32447 standard; cDNA; 1053 BP.

XX

AC AAF32447;

XX

DT 18-APR-2001 (first entry)

XX

DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:12.

XX

KW Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.

XX

OS Homo sapiens.

XX

FN WO200100799-A1.

XX

PD 04-JAN-2001.

XX

PF 27-JUN-2000; 2000WO-JP004189.

XX

PR 28-JUN-1999; 99JP-00181131.

PR 06-JUL-1999; 99JP-00192391.

PR 21-JAN-2000; 2000JP-00017770.

XX

XX (TAKE) TAKEDA CHEM IND LTD.

PA (NAKA/) NAKAMURA Y.

PA

PI Nakamura Y, Arakawa H, Tanaka H;

XX

XX WPI; 2001-112446/12.

XX

PT Ribonucleotide reductase involved in DNA repair and DNA encoding it, for

XX diagnosis, treatment and prevention of cancer.

XX

PS Claim 6; Page 96-97; 102pp; Japanese.

XX

CC The present invention describes a human ribonucleotide reductase

CC designated TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of

CC the DNA repair mechanism and its activity is induced by p53. It can be

CC used for the treatment, prevention and diagnosis of a wide range of

CC cancers. The present sequence represents a human ribonucleotide reductase

CC related sequence which is given in the exemplification of the present

CC invention

Adk70302 Respirato
Adj56536 Murine cD
Aah73325 Human cer
Aas44745 Human ful
Aav05641 Human rib
Ado57308 DNA enco
Aas41006 cDNA enco
Ab190228 Human pol
Aah07707 Human cDN
Aas79474 DNA enco
Abl14627 Drosophil
Aas62387 cDNA sequ
Aas51874 Human sec
Abt20895 Aspergill
Aas4544 Swinepox
Aas4949 Reverse c
Aaf84948 Nucleotid
Ab114626 Drosophil
AdS58378 Bacterial
Abt19075 Aspergill
Adt47814 Bacterial
AbS32250 Candida a
AdS49379 Bacterial
Acc61238 Gene sequ
Adk63241 Disease t

21 578.4 54.9 2215 12 ADK70302
22 578.4 54.9 2216 10 ADJ56536
23 578.4 54.9 3294 4 AAH73325
24 569 54.0 2641 4 AAS44745
25 568.8 54.0 1170 2 AAV05641
26 562.6 53.4 1328 12 ADO57308
27 547.8 52.0 978 4 AAS41006
28 545.4 51.8 977 6 ABL90228
29 527.2 50.1 608 4 AAH07707
30 514 48.8 1371 5 AAS79474
31 463 44.0 1289 4 ABL14627
32 442 42.0 1146 6 AAS62387
33 439.4 41.7 481 2 AAS51874
34 433.8 41.2 1218 8 ABT20895
35 429.8 40.8 14176 2 AAT84564
36 429.8 40.8 14176 4 AAF84949
37 429.8 40.8 14176 4 AAF84948
38 420.4 39.9 3945 4 ABL14626
39 404.8 38.4 963 13 ADS58378
40 398.5 37.9 1146 8 ABT19075
41 394.2 37.4 1173 13 ADT47814
42 393.6 37.4 1242 6 ABZ32250
43 388.2 36.9 1292 13 ADS49379
44 382 36.3 1200 10 ACC61238
45 382 36.3 1200 10 ADK63241

XX SQ Sequence 1053 BP; 314 A; 191 C; 246 G; 302 T; 0 U; 0 Other;

Query Match 100.0%; Score 1053; DB 4; Length 1053;
Best Local Similarity 100.0%; Pred. No. 9e-290;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGACCCCGAAGCGCGGCGGCTGGATCAGAGTCAGAGATCATCTTCA 60
DB 1 ATGGCGACCCCGAAGCGCGGCGGCTGGATCAGAGTCAGAGATCATCTTCA 60
QY 61 GACACCAACGAAAGTGAATAAAGTCAAATGAAGAGCCACTCTCTAAGAAAGTCTCGC 120
DB 61 GACACCAACGAAAGTGAATAAAGTCAAATGAAGAGCCACTCTCTAAGAAAGTCTCGC 120
QY 121 CGGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAATGTATATAACAGGCACAG 180
DB 121 CGGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAATGTATATAACAGGCACAG 180
QY 181 GCTTCCTCTGACAGCAGAGAGGTGTACTTATCAAGGATCTCCCTCACTGGAAACAAG 240
DB 181 GCTTCCTCTGACAGCAGAGAGGTGTACTTATCAAGGATCTCCCTCACTGGAAACAAG 240
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTTACATCTTACGCTTTTTCGAGCCAGTGAT 300
DB 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTTACATCTTACGCTTTTTCGAGCCAGTGAT 300
QY 301 GGAATGTAATGAAATTTGTGAGCGCTTTAGTCAGGAGGTGCAGGTCAGAGGGCT 360
DB 301 GGAATGTAATGAAATTTGTGAGCGCTTTAGTCAGGAGGTGCAGGTCAGAGGGCT 360
QY 361 CGCTGTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
DB 361 CGCTGTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGTAA 480
DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGTAA 480
QY 481 ACCATGCCCTATGTTAAGAAAGAGAGATTGGGCTTCGATGAGATAGAGAGAGAA 540
DB 481 ACCATGCCCTATGTTAAGAAAGAGAGATTGGGCTTCGATGAGATAGAGAGAGAA 540
QY 541 TCTACTTTTGGGAAGAGTGTGGCTTTGCTGTAGAGAGGATTTCTTCACGA 600
DB 541 TCTACTTTTGGGAAGAGTGTGGCTTTGCTGTAGAGAGGATTTCTTCACGA 600
QY 601 TCTTTTGTCTATATTTCTGGCTAAGAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC 660
DB 601 TCTTTTGTCTATATTTCTGGCTAAGAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC 660
QY 661 AATGAACATCAGCAGAGATGAAGACTTCTCTGTGACTTTGCTGCTGATGTTCCAA 720
DB 661 AATGAACATCAGCAGAGATGAAGACTTCTCTGTGACTTTGCTGCTGATGTTCCAA 720
QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTGCAGGAGATCATTTGTCATGCTGCA 780
DB 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTGCAGGAGATCATTTGTCATGCTGCA 780
QY 781 ATTGACGAGGAGTTTAAACAGAAAGCCTTGGCAGTTGGCTCATTTGGAATGAATTCATT 840
DB 781 ATTGACGAGGAGTTTAAACAGAAAGCCTTGGCAGTTGGCTCATTTGGAATGAATTCATT 840
QY 841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTAATTTGGAACCTTGGATTTCA 900
DB 841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTAATTTGGAACCTTGGATTTCA 900
QY 901 AAGGTTTTTCAGGCAGAAATCCTTTTGAATTTTGAATTTGAATTTTGAAGAGAAA 960
DB 901 AAGGTTTTTCAGGCAGAAATCCTTTTGAATTTTGAATTTTGAAGAGAGAAA 960
QY 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAC 1020
DB 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAC 1020

DB 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAC 1020
QY 1021 ACAGATAACGCTCTTCACTTGGATGCAGATTTT 1053
DB 1021 ACAGATAACGCTCTTCACTTGGATGCAGATTTT 1053

RESULT 2
AAF32440
ID AAF32440 standard; cDNA; 1081 BP.
XX AAF32440;
AC XX
DT 18-APR-2001 (first entry)
XX Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:4.
DE Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
KW Homo sapiens.
OS OS
PN WO200100799-A1.
XX
PD 04-JAN-2001.
XX
PF 27-JUN-2000; 2000WO-JP004189.
XX
PR 28-JUN-1999; 99JP-00181131.
PR 06-JUL-1999; 99JP-00192391.
PR 21-JAN-2000; 2000JP-00017770.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
PA (NAKA/) NAKAMURA Y.
XX Nakamura Y, Arakawa H, Tanaka H;
PI WPI; 2001-112446/12.
XX
DR Ribonucleotide reductase involved in DNA repair and DNA encoding it, for diagnosis, treatment and prevention of cancer.
XX
PT Example 2; Page 91; 102pp; Japanese.
XX
CC The present invention describes a human ribonucleotide reductase designated TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of the DNA repair mechanism and its activity is induced by p53. It can be used for the treatment, prevention and diagnosis of a wide range of cancers. The present sequence represents a human ribonucleotide reductase related sequence which is used in an example from the present invention
CC
XX Sequence 1081 BP; 322 A; 199 C; 252 G; 308 T; 0 U; 0 Other;

Query Match 100.0%; Score 1053; DB 4; Length 1081;
Best Local Similarity 100.0%; Pred. No. 9.1e-290;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGAAAGCGCGGAGCGCGGCTGGATCAGAGTCAGAGATCATCTTCA 60
DB 20 ATGGGCGACCCGAAAGCGCGGAGCGCGGCTGGATCAGAGTCAGATCATCTTCA 79
QY 61 GACACCAACGAAAGTGAATAAAGTCAAATGAAGAGCCACTCTCTAAGAAAGTCTCGC 120
DB 80 GACACCAACGAAAGTGAATAAAGTCAAATGAAGAGCCACTCTCTAAGAAAGTCTCGC 139
QY 121 CGGTTTGTCTATCTTCCATCCAGTACCTCGATATTTGGAAATGTATATAACAGGCACAG 180
DB 140 CGGTTTGTCTATCTTCCATCCAGTACCTCGATATTTGGAAATGTATATAACAGGCACAG 199
QY 181 GCTTCCTCTGAGCAGCAGAGAGGTGTACTTATCAAGAGATCTCCCTCACTGGAAACAAG 240
DB 200 GCTTCCTCTGAGCAGCAGAGAGGTGTACTTATCAAGAGATCTCCCTCACTGGAAACAAG 259
QY 241 CTTAAGACGATGAGAGTACTTCTCATCTTACATCTTTCAGCCTTTTTCGAGCCAGTGAT 300


```
Db 260 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCATCTAGCCCTTTTTCAGCGAGTGAT 319
Qy 301 GGAAATTTGTAATAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGAGGTTCCAGAGGCT 360
Db 320 GGAAATTTGTAATAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGAGGTTCCAGAGGCT 379
Qy 361 CGCTGTTTCTATGGCTTTCAAAATTTCTATCGAGAAATGTTTCACTCAGAGATGTACAGTTTG 420
Db 380 CGCTGTTTCTATGGCTTTCAAAATTTCTATCGAGAAATGTTTCACTCAGAGATGTACAGTTTG 439
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATGAA 480
Db 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATGAA 499
Qy 481 ACCATGCCCTATGTTAAGAAAGAGAGATGGGCTTGGGATGGATAGCAGATAGAAA 540
Db 500 ACCATGCCCTATGTTAAGAAAGAGAGATGGGCTTGGGATGGATAGCAGATAGAAA 559
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTCTTCTTCAGGA 600
Db 560 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTCTTCTTCAGGA 619
Qy 601 TCTTTTGTGCTATATTTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
Db 620 TCTTTTGTGCTATATTTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTCC 679
Qy 661 AATGAACCTCATCAGCAGAGATGAGAGCTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 720
Db 680 AATGAACCTCATCAGCAGAGATGAGAGCTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 739
Qy 721 TACTTAGTAAATAAAGCTTTCAGAGAAAGGCTCAGGAGATCACTTGTGATGCTGTCAA 780
Db 740 TACTTAGTAAATAAAGCTTTCAGAGAAAGGCTCAGGAGATCACTTGTGATGCTGTCAA 799
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGGCCAGTGGCTCATTTGGAATGAATTCATT 840
Db 800 ATTGAGCAGGAGTTTAAACAGAGCCTTGGCCAGTGGCTCATTTGGAATGAATTCATT 859
Qy 841 TTGATGAAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTTCTCA 900
Db 860 TTGATGAAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTTCTCA 919
Qy 901 AAGGTTTTTTCAGCAGAAATCCTTTTGATTTTATGGAACAACTTTTAGAAGAAA 960
Db 920 AAGGTTTTTTCAGCAGAAATCCTTTTGATTTTATGGAACAACTTTTAGAAGAAA 979
Qy 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGATATCAGCGTTTTTGCAGTTATGGCAGAAAC 1020
Db 980 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGATATCAGCGTTTTTGCAGTTATGGCAGAAAC 1039
Qy 1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053
Db 1040 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1072
```

RESULT 3

AAF32438
ID AAF32438 standard; cDNA; 1053 BP.

XX AAF32438;

AC AAF32438;

XX 18-APR-2001 (first entry)

DE Human ribonucleotide reductase TP53R2H nucleotide sequence SEQ ID NO:2.

XX Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.

OS Homo sapiens.

XX WO200100799-A1.

XX 04-JAN-2001.

PD

```
XX 27-JUN-2000; 2000WO-JP004189.
XX
PR 28-JUN-1999; 99JP-00181131.
PR 06-JUL-1999; 99JP-00192391.
PR 21-JAN-2000; 2000JP-00017770.
XX
PA (TAKE ) TAKEDA CHEM IND LTD.
PA (NAKA/) NAKAMURA Y.
XX
PI Nakamura Y, Arakawa H, Tanaka H;
XX
DR WPI; 2001-112446/12.
DR P-PSDB; AAB69050.
XX
PT Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
PT diagnosis, treatment and prevention of cancer.
XX
PS Claim 5; Fig 1-3; 102pp; Japanese.
XX
CC The present sequence encodes a human ribonucleotide reductase designated
CC TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of the DNA
CC repair mechanism and its activity is induced by p53. It can be used for
CC the treatment, prevention and diagnosis of a wide range of cancers
XX
SQ Sequence 1053 BP; 314 A; 192 C; 246 G; 301 T; 0 U; 0 Other; ..
Query Match 99.8%; Score 1051.4; DB 4; Length 1053;
Best Local Similarity 99.9%; Pred. No. 2.6e-289;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGGGCGACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
Db 1 ATGGGCGACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
Qy 61 GACACCAACGAAGTGAATTAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAGTGAATTAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Qy 121 CGGTTGTGCTATCTTCCAAATCCAGTACCTGATATTTGGAAAATGTATAAACAGGCACAG 180
Db 121 CGGTTGTGCTATCTTCCAAATCCAGTACCTGATATTTGGAAAATGTATAAACAGGCACAG 180
Qy 181 GCTTCTCTTGGACAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAAG 240
Db 181 GCTTCTCTTGGACAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAAG 240
Qy 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCCACATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCCACATCTTAGCCCTTTTTCAGCCAGTGAT 300
Qy 301 GGAATTTGTAATAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGAGGTTCCAGAGGCT 360
Db 301 GGAATTTGTAATAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGAGGTTCCAGAGGCT 360
Qy 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAAGATGTTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAAGATGTTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAGAGAGATGGGCTTGGGATGGATAGCAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAGAGAGATGGGCTTGGGATGGATAGCAGATAGAAA 540
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTCTTCTTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTCTTCTTCAGGA 600
Qy 601 TCTTTTGTGCTATATTTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
```

Db 601 TCTTTTGTGCTATATTCTTGGCTAAAGAGAGAGGCTTTATGCCAGGACTCACTTTTTC 660
 QY 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
 Db 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
 QY 721 TACTTTAGTAATAAAGCCTTCAGAAGAAAGGCTCAGGAGATCATTTTGTGATGCTGTCAA 780
 Db 721 TACTTTAGTAATAAAGCCTTCAGAAGAAAGGCTCAGGAGATCATTTTGTGATGCTGTCAA 780
 QY 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATGCAAT 840
 Db 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATGCAAT 840
 QY 841 TTGATGAACACTACATTCAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTCCTCA 900
 Db 841 TTGATGAACACTACATTCAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTCCTCA 900
 QY 901 AAGGTTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGGAAAA 960
 Db 901 AAGGTTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGGAAAA 960
 QY 961 ACAAAATTTCTTGAGAAACAGATTTTCAGAGTATCAGGTTTTCAGTTTTCAGTATGCGAGAAAC 1020
 Db 961 ACAAAATTTCTTGAGAAACAGATTTTCAGAGTATCAGGTTTTCAGTTTTCAGTATGCGAGAAAC 1020
 QY 1021 ACAGATAAGCTCTTACCTTGGATGCGAGATTTT 1053
 Db 1021 ACAGATAAGCTCTTACCTTGGATGCGAGATTTT 1053

RESULT 4

AAAL2411
 ID AAAL2411 standard; cDNA; 2596 BP.
 AC AAAL2411;
 DT 25-JUL-2000 (first entry)
 DE cDNA encoding a human RNA-associated protein.

Human; RNA-associated protein; cell proliferation; cancer; inflammation;
 immune response; reproductive disorder; actinic keratosis;
 KX atherosclerosis; arteriosclerosis; bursitis; cirrhosis; hepatitis;
 KW mixed connective tissue disease; myelofibrosis; primary thrombocythemia;
 KW paroxysmal nocturnal hemoglobinuria; polycythemia vera; psoriasis;
 KW trauma; ss.

OS Homo sapiens.

Key Location/Qualifiers
 CDS 56..1111
 FT /*tag= a
 FT /product= "RNA-associated protein"

PN WO200015799-A2.

XX 23-MAR-2000.

PF 17-SEP-1999; 99WO-US021688.

PR 17-SEP-1998; 98US-00156039.

PR 22-SEP-1998; 98US-00158720.

PR 04-NOV-1998; 98US-00186815.

PR 08-APR-1999; 99US-0128660P.

XX (INCY-) INCYTE PHARM INC.

PI Tang YT, Corley NC, Guegler KJ, Gorgone GA, Patterson C;

PI Hillman JL, Baughn MR, Lal P, Azimzai Y, Yue H, Yang J;

XX WPI; 2000-271437/23.

DR P-PSDB; AAY84439.

XX New polypeptides and polynucleotides, useful for preventing and treating
 PT a disorder associated with increased or decreased expression of RNA
 PT associated proteins.
 PS Claim 9; Page 120-121; 131pp; English.
 XX
 CC The present sequence encodes a human RNA-associated protein. The
 CC expression of RNA-associated proteins is closely associated with
 CC reproductive tissues, nervous tissues, cell proliferation including
 CC cancer, inflammation and immune responses, and so they may be used for
 CC diagnosis, treatment or prevention of cell proliferative,
 CC immune/inflammatory disorders, and reproductive disorders. Diseases and
 CC disorders which may be treated include actinic keratosis,
 CC atherosclerosis, arteriosclerosis, bursitis, cirrhosis, hepatitis, mixed
 CC connective tissue disease, myelofibrosis, paroxysmal nocturnal
 CC hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia
 CC and cancers, and trauma
 XX
 SQ Sequence 2596 BP; 783 A; 418 C; 510 G; 885 T; 0 U; 0 Other;

Query Match 99.8%; Score 1051.4; DB 3; Length 2596;
 Best Local Similarity 99.9%; Pred. No. 3.9e-289;
 Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGGCGCACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
 Db 56 ATGGCGCACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 115
 QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCCTAAGAAAAGATTTCTCG 120
 Db 116 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCCTAAGAAAAGATTTCTCG 175
 QY 121 CGGTTTGTCTATCTTTTCCAATCCAGTACCTCTGATATTTGGAAAAATGTATAAAGGACAG 180
 Db 176 CGGTTTGTCTATCTTTTCCAATCCAGTACCTCTGATATTTGGAAAAATGTATAAAGGACAG 235
 QY 181 GCTTCTCTTGGCAGCAGAGAGGTTGACTTATCAAAAGGATCTCCCTCAGTGGAAACAG 240
 Db 236 GCTTCTCTTGGCAGCAGAGAGGTTGACTTATCAAAAGGATCTCCCTCAGTGGAAACAG 295
 QY 241 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGTAT 300
 Db 296 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGTAT 355
 QY 301 GGAATTGTAATGAAAATTTGGTGGAGCGCTTTAGTTCAGGAGGTCAGAGTTCAGAGGCT 360
 Db 356 GGAATTGTAATGAAAATTTGGTGGAGCGCTTTAGTTCAGGAGGTCAGAGTTCAGAGGCT 415
 QY 361 CGCTGTTTCTATGCTTTCAAAATTTCTCATCAGAGATGTTCACTCAGAGATGTACAGTTTG 420
 Db 416 CGCTGTTTCTATGCTTTCAAAATTTCTCATCAGAGATGTTTCATCTCAGAGATGTACAGTTTG 475
 QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATCAATTTGAA 480
 Db 476 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATCAATTTGAA 535
 QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTTTCGATGGATAGCAGATGAAAA 540
 Db 536 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTTTCGATGGATAGCAGATGAAAA 595
 QY 541 TCTACTTTTGGGAAAAGAGTGGGCTTTTCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
 Db 596 TCTACTTTTGGGAAAAGAGTGGGCTTTTCTGCTGTAGAGGAGTTTCTTCTCAGGA 655
 QY 601 TCTTTTGTCTGTATTTCTGCTTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
 Db 656 TCTTTTGTCTGTATTTCTGCTTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 715
 QY 661 AATGAACATCATCAGCAGAGATGAAGGACTTTCACCTGTGACTTTTGTCTGCTGATGTTCCAA 720
 Db 716 AATGAACATCATCAGCAGAGATGAAGGACTTTCACCTGTGACTTTTGTCTGCTGATGTTCCAA 775

```
QY 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGTCAAA 780
Db 776 TACTTAGTAAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGTCAAA 835
QY 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATTGCATT 840
Db 836 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATTGCATT 895
QY 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGTGAACTTTGATTTCTCA 900
Db 896 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGTGAACTTTGATTTCTCA 955
QY 901 AAGGTTTTTTCAGGACAGAAATCCTTTGATTTTATGAAAAACATTTCTTTAGAAAGGAAAA 960
Db 956 AAGGTTTTTTCAGGACAGAAATCCTTTGATTTTATGAAAAACATTTCTTTAGAAAGGAAAA 1015
QY 961 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1020
Db 1016 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1075
QY 1021 ACAGATAACGCTCTTACCTTTGGATGAGATTTT 1053
Db 1076 ACAGATAACGCTCTTACCTTTGGATGAGATTTT 1108

RESULT 5
AAF32439
ID AAF32439 standard; cDNA; 4955 BP.
XX
AC AAF32439;
XX
DT 18-APR-2001 (first entry)
XX
DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:3.
XX
KW Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
XX
OS Homo sapiens.
XX
PN WO200100799-A1.
XX
PD 04-JAN-2001.
XX
PF 27-JUN-2000; 2000WO-JP004189.
XX
PR 28-JUN-1999; 95JP-00181131.
PR 06-JUL-1999; 95JP-00192391.
PR 21-JAN-2000; 2000JP-00017770.
XX
PA (TAKE ) TAKEDA CHEM IND LTD.
PA (NAKA/) NAKAMURA Y.
XX
PI Nakamura Y, Arakawa H, Tanaka H;
XX
DR WPI; 2001-112446/12.
XX
PT Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
PT diagnosis, treatment and prevention of cancer.
XX
PS Example 2; Page 87-90; 102pp; Japanese.
XX
CC The present invention describes a human ribonucleotide reductase
CC designated TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of
CC the DNA repair mechanism and its activity is induced by p53. It can be
CC used for the treatment, prevention and diagnosis of a wide range of
CC cancers. The present sequence represents a human ribonucleotide reductase
CC related sequence which is used in an example from the present invention
XX
SQ Sequence 4955 BP; 1526 A; 803 C; 999 G; 1627 T; 0 U; 0 Other;
Query Match 99.8%; Score 1051.4; DB 4; Length 4955;
Best Local Similarity 99.9%; Pred. No. 5.2e-289;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 ATGGCGACCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 245 ATGGCGACCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCATCTCTAAGAAAGAGTTCTCGC 120
Db 305 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCATCTCTAAGAAAGAGTTCTCGC 364
QY 121 CGGTTTGTCTATCTTTCAATCCAGTACCTGTATATTTGGAATAATGTATAACAGGCACAG 180
Db 365 CGGTTTGTCTATCTTTCAATCCAGTACCTGTATATTTGGAATAATGTATAACAGGCACAG 424
QY 181 GCTTCTCTTCTGACAGCAGAGAGGTTGACTTATCAAAAGAGTCTCCCTCACTTGAACCAAG 240
Db 425 GCTTCTCTTCTGACAGCAGAGAGGTTGACTTATCAAAAGAGTCTCCCTCACTTGAACCAAG 484
QY 241 CTTAAAGCAGATGAGAGTACTTTCATCTCTCACATCTTAGCCTTTTTCAGCAGCCAGTAT 300
Db 485 CTTAAAGCAGATGAGAGTACTTTCATCTCTCACATCTTAGCCTTTTTCAGCAGCCAGTAT 544
QY 301 GGAATTTGTAATGAAATTTTGGTGGAGCGCTTTTAGTCAGAGAGTTCAGAGGCT 360
Db 545 GGAATTTGTAATGAAATTTTGGTGGAGCGCTTTTAGTCAGAGAGTTCAGAGGCT 604
QY 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGAATGTTTCACTCAGAGATCTACAGTTTG 420
Db 605 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGAATGTTTCACTCAGAGATCTACAGTTTG 664
QY 421 CTGATAGACACTTATACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTGAA 480
Db 665 CTGATAGACACTTATACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTTCGATGATGATAGAGATAGAAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTTCGATGATGATAGAGATAGAAAA 784
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGAGTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGAGTTTCTTCTCAGGA 844
QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGGTTCTTATGCCAGAGTCTTATGCTGCTTCC 660
Db 845 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGGTTCTTATGCCAGAGTCTTATGCTGCTTCC 904
QY 661 AATGAACTCATCAGCAGAGATGAAGGACTTTCACCTGTGACTTTGCTGTGCTGATGTTCAA 720
Db 905 AATGAACTCATCAGCAGAGATGAAGGACTTTCACCTGTGACTTTGCTGTGCTGATGTTCAA 964
QY 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGTCAAA 780
Db 965 TACTTAGTAAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGTCAAA 1024
QY 781 ATTGAGCAGGAGTTTAAACAGAGCCTTCCAGTTGGCCTCATTTGGAATGAATTGCATT 840
Db 1025 ATTGAGCAGGAGTTTAAACAGAGCCTTCCAGTTGGCCTCATTTGGAATGAATTGCATT 1084
QY 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGGAACATTTGGATTCTCA 900
Db 1085 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGGAACATTTGGATTCTCA 1144
QY 901 AAGGTTTTTTCAGGACAGAAATCCTTTGATTTTATGAAAAACATTTCTTTAGAAAGGAAAA 960
Db 1145 AAGGTTTTTTCAGGACAGAAATCCTTTGATTTTATGAAAAACATTTCTTTAGAAAGGAAAA 1204
QY 961 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1020
Db 1205 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1264
QY 1021 ACAGATAACGCTCTTACCTTTGGATGAGATTTT 1053
Db 1265 ACAGATAACGCTCTTACCTTTGGATGAGATTTT 1297
```

RESULT 6
 ID ADR24210 standard; DNA; 4955 BP.
 AC ADR24210;
 DT 21-OCT-2004 (first entry)
 XX Breast cancer prognosis marker #71.
 XX ds; breast cancer; prognosis; gene expression; diagnosis.
 KW Homo sapiens.
 OS WO2004065545-A2.
 PN 05-AUG-2004.
 XX 15-JAN-2004; 2004WO-US001100.
 XX 15-JAN-2003; 2003US-00342887.
 XX (ROSE-) ROSETTA INPHARMATICS LLC.
 PA (NECA-) NETHERLANDS CANCER INST.
 XX Van't Veer LJ, He Y;
 XX WPI; 2004-593473/57.
 DR
 CC The invention relates to a method of classifying a breast cancer patient
 CC according to prognosis by determining the similarity between the level of
 CC expression of each of five genes for which markers are listed in the
 CC specification, in a cell sample taken from the breast cancer patient, to
 CC control levels of expression for each respective five genes to obtain a
 CC patient similarity value. The methods are useful for classifying a breast
 CC cancer patient according to prognosis. Kits and computer program products
 CC are useful for data analysis using the diagnostic, prognostic and
 CC statistical methods of the invention. This sequence corresponds to a
 CC marker used in the method of the invention.
 XX
 SQ Sequence 4955 BP; 1526 A; 803 C; 999 G; 1627 T; 0 U; 0 Other;
 Query Match 99.8%; Score 1051.4; DB 13; Length 4955;
 Best Local Similarity 99.9%; Pred. NO. 5.2e-289;
 Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGGGCGACCCGGAAAGCGCGGAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
 DB |||||
 QY 245 ATGGGCGACCCGGAAAGCGCGGAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 304
 DB |||||
 QY 61 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTGCG 120
 DB |||||
 QY 305 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTGCG 364
 DB |||||
 QY 121 CGGTTTGTCATCTTCCCAATCCAGTACCTGATATTTGGAATATGTAAACAGGCACAG 180
 DB |||||
 QY 365 CGGTTTGTCATCTTCCCAATCCAGTACCTGATATTTGGAATATGTAAACAGGCACAG 424
 DB |||||
 QY 181 GCTTCTCTTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
 DB |||||
 QY 425 GCTTCTCTTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 484
 DB |||||
 QY 241 CTTAAAGCAGATGAGAAGTACTTCTCTCTCATCTTACATCTTTCAGCCCTTTTTCAGCCAGTGAT 300
 DB |||||
 QY 485 CTTAAAGCAGATGAGAAGTACTTCTCTCTCATCTTACATCTTTCAGCCCTTTTTCAGCCAGTGAT 544
 DB |||||

QY 301 GGAATTTAAATGAAATTTTGGTGGAGCGCTTTAGTCAAGAGGTGAGGTTCCAGAGGCT 360
 DB |||||
 QY 545 GGAATTTGTAATGAAATTTTGGTGGAGCGCTTTAGTCAAGAGGTGAGGTTCCAGAGGCT 604
 DB |||||
 QY 361 CGCTGTTTCTATGCTTTCAAAATTTCTCATCGAATGTTCACTCAGAGATGTACAGTTTG 420
 DB |||||
 QY 605 CGCTGTTTCTATGCTTTCAAAATTTCTCATCGAATGTTCACTCAGAGATGTACAGTTTG 664
 DB |||||
 QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAAATCAATTTGAA 480
 DB |||||
 QY 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAAATCAATTTGAA 724
 DB |||||
 QY 481 ACCATGCCCTATGTTAAAGAAAAAGCAGATTGGGCTTTGCGATGATGATGATGATGATG 540
 DB |||||
 QY 725 ACCATGCCCTATGTTAAAGAAAAAGCAGATTGGGCTTTGCGATGATGATGATGATGATG 784
 DB |||||
 QY 541 TCTACTTTTGGGAAAGAGTGGCTTTGCTGCTGTAGAGAGGATTTTCTTCTCAGGA 600
 DB |||||
 QY 785 TCTACTTTTGGGAAAGAGTGGCTTTGCTGCTGTAGAGAGGATTTTCTTCTCAGGA 844
 DB |||||
 QY 601 TCTTTTGTCTATATTTCTGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
 DB |||||
 QY 845 TCTTTTGTCTATATTTCTGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 904
 DB |||||
 QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGATCTTGTGCTGATGTTCCAA 720
 DB |||||
 QY 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGATCTTGTGCTGATGTTCCAA 964
 DB |||||
 QY 721 TACTTAGTAATAAGCCTTCAGAAAGAGGTCAAGGATCATTTGTTGATGCTGTCAAA 780
 DB |||||
 QY 965 TACTTAGTAATAAGCCTTCAGAAAGAGGTCAAGGATCATTTGTTGATGCTGTCAAA 1024
 DB |||||
 QY 781 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGCAGTGGCCTCATTTGGAATGAATTCATT 840
 DB |||||
 QY 1025 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGCAGTGGCCTCATTTGGAATGAATTCATT 1084
 DB |||||
 QY 841 TTGATGAAACAGTACATTTAGCTTTGAGTCAAGATTAATTTGTAATCTGATCTCA 900
 DB |||||
 QY 1085 TTGATGAAACAGTACATTTAGCTTTGAGTCAAGATTAATTTGTAATCTGATCTCA 1144
 DB |||||
 QY 901 AAGGTTTTTCAAGCAGAAAAATCTTTTGAATTTTATGGAACATTTCTTTAGAGAAAA 960
 DB |||||
 QY 1145 AAGGTTTTTCAAGCAGAAAAATCTTTTGAATTTTATGGAACATTTCTTTAGAGAAAA 1204
 DB |||||
 QY 961 ACAAATTTCTTTGAAACAGGTTTTCAGATATCAGCGTTTTCAGTTTATGCGAGAACCC 1020
 DB |||||
 QY 1205 ACAAATTTCTTTGAAACAGGTTTTCAGATATCAGCGTTTTCAGTTTATGCGAGAACCC 1264
 DB |||||
 QY 1021 ACAGATAACGTTCTTCACTTGGATGATGATTTT 1053
 DB |||||
 QY 1265 ACAGATAACGTTCTTCACTTGGATGATGATTTT 1297
 DB |||||
 RESULT 7
 AAH14924
 ID AAH14924 standard; cDNA; 1601 BP.
 XX AC AAH14924;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA sequence SEQ ID NO:12810.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 OS Homo sapiens.
 XX EF1074617-A2.
 PN 07-FEB-2001.
 XX
 XX 28-JUL-2000; 2000EP-00116126.
 XX

29-JUL-1999; 95JP-00248036.
 27-AUG-1999; 95JP-00300253.
 11-JAN-2000; 2000JP-00118776.
 02-MAY-2000; 2000JP-00183767.
 09-JUN-2000; 2000JP-00241899.
 (HELI-) HELIX RES INST.
 PA
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR WPI; 2001-318749/34.
 XX
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 PS Claim 8; SEQ ID NO 12810; 2537pp + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dr primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH013628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 XX
 SQ Sequence 1601 BP; 484 A; 290 C; 348 G; 479 T; 0 U; 0 Other;
 Query Match 99.5%; Score 1048.2; DB 4; Length 1601;
 Best Local Similarity 99.7%; Pred. No. 2.5e-288;
 Matches 1050; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 1 ATGGCGGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATGAGATCACTTCA 60
 40 ATGGCGGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATGAGATCACTTCA 99
 61 GACACCAACGAAAGTGAATGAATGAATGAATGAATGAATGAATGAATGAAT 120
 100 GACACCAACGAAAGTGAATGAATGAATGAATGAATGAATGAATGAATGAAT 159
 121 CGTTTGTTCATCTTTCCATCCAGTACCTGATATTTGGAAATGTATAAAGGACAG 180
 160 CGTTTGTTCATCTTTCCATCCAGTACCTGATATTTGGAAATGTATAAAGGACAG 219
 181 GCTTCCTTCTGGACAGGAGAGGTTGACTATCAAGATCTCCCTCACTGGAAACAG 240
 220 GCTTCCTTCTGGACAGGAGAGGTTGACTATCAAGATCTCCCTCACTGGAAACAG 279
 241 CTTAAGCAGATGAGAAGTACTTCTCTCATCTCTAGCCCTTTTTCAGCCAGTAT 300
 280 CTTAAGCAGATGAGAAGTACTTCTCTCATCTCTAGCCCTTTTTCAGCCAGTAT 339
 301 GGAATTTGTAATGAAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTCCAGAGGCT 360
 340 GGAATTTGTAATGAAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTCCAGAGGCT 399

361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 420
 400 CGCTGCTTCTATGGCTTTCAAATTTCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 459
 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA 480
 460 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA 519
 481 ACCATGCCCTATGTTAAGAAAAGGAGATGGGCTTTGCCGATGATAGCAGATAGAAA 540
 520 ACCATGCCCTATGTTAAGAAAAGGAGATGGGCTTTGCCGATGATAGCAGATAGAAA 579
 541 TCTACTTTTGGGGAAGAGTGGTGGCCCTTCTGCTGTAGAGAGGATTTTCTTCTCAGAG 600
 580 TCTACTTTTGGGGAAGAGTGGTGGCCCTTCTGCTGTAGAGAGGATTTTCTTCTCAGAG 639
 601 TCTTTTGTCTCTATTAATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
 640 TCTTTTGTCTCTATTAATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC 699
 661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 720
 700 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 759
 721 TACTTAGTAATAAGCCTTTCAGAAAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
 760 TACTTAGTAATAAGCCTTTCAGAAAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 819
 781 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGCAGTGGCCTCATTTGGAATGAATTCATTC 840
 820 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGCAGTGGCCTCATTTGGAATGAATTCATTC 879
 841 TTGATGAACAGTACATGATGATGTTGTAGTGCAGAGATTAATTTGGAATTCATTCATTC 900
 880 TTGATGAACAGTACATGATGATGTTGTAGTGCAGAGATTAATTTGGAATTCATTCATTC 939
 901 AAGTTTTTTCAGCAGAGAAATCCTTTGATTTTATGAAAACATTTCTTTAGAGGAAA 960
 940 AAGTTTTTTCAGCAGAGAAATCCTTTGATTTTATGAAAACATTTCTTTAGAGGAAA 999
 961 ACAAATTTCTTTCAGAAACAGATTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAGAAC 1020
 1000 ACAAATTTCTTTCAGAAACAGATTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAGAAC 1059
 1021 ACAGATAACGCTTTCACCTTGGATGCGAGATTTT 1053
 1060 ACAGATAACGCTTTCACCTTGGATGCGAGATTTT 1092
 RESULT 8
 ADQ57092
 ID ADQ57092 standard; DNA; 706 BP.
 XX
 AC ADQ57092;
 XX
 XX
 DT 21-OCT-2004 (first entry)
 XX
 DE Novel canine microarray-related DNA sequence SeqID8394.
 XX
 XX canine microarray; drug screening; toxicity assay;
 KW environmental pollutant; cellular response; gene expression profile;
 KW toxic response; liver necrosis; fatty liver disease;
 KW protein adduct formation; hepatitis; dog; ds.
 XX
 OS Canis familiaris.
 XX
 XX
 PN WO2004063324-A2.
 XX
 PD 29-JUL-2004.
 XX
 XX 05-MAY-2003; 2003WO-US013853.
 XX
 XX 03-MAY-2002; 2002US-0377240P.
 PR

PT neurological, inflammatory disorders and for use in arrays for detection.

PS Claim 1; SEQ ID NO 514; 153pp; English.

XX
XX
CC Sequences AAS44576-AAS44919 represent full-length polynucleotides and
CC contig polynucleotides encoding polypeptides of the invention. The DNA
CC and protein sequences are useful for the treatment, diagnosis and
CC prevention of various types of disorder in a mammalian subject such as a
CC human, dog, monkey, mouse, hamster or rat. The disorders include cancers
CC such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such
CC as multiple sclerosis, connective tissue disease, rheumatoid arthritis,
CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's
CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
CC bowel disease. The sequences exhibit activity relating to angiogenesis,
CC cell proliferation, cell differentiation, stem cell growth factor,
CC activin or inhibin. Therefore, they can be used to manipulate stem cells
CC in culture to give rise to neuroepithelial cells that can be used to
CC augment or replace cells damaged by illness, accidental damage or genetic
CC disorders. The sequences may also be used for regeneration of bone,
CC cartilage, tendons and ligaments and in tissue repair and burn healing.
CC Note: Some sequences for this patent did not form part of the printed
CC specification, but were obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 2482 BP; 707 A; 598 C; 522 G; 655 T; 0 U; 0 Other;

Query Match 55.1%; Score 580; DB 4; Length 2482;

Best Local Similarity 75.1%; Pred. No. 1.4e-154;

Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY	89	ATGAAGAGCCATCTCTAAGAAAGATTCTCGCGGTTTGTCATCTTCCCAATCCAGTACC	148
DB	2086	AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTGTGATCTTCCCATCGAGTACC	2027
QY	149	CTGATATTGGAAATGTAACAGGCACAGGCTTCCTTGGACAGCAGAGAGGTG	208
DB	2026	ATGATATCTGGCAGATGTAAGAGGCGAGAGGCTTCCTTTGGACCGCGAGGAGGTG	1967
QY	209	ACTTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAGTCTTCATCT	268
DB	1966	ACCTCTCAGGACATTCAGCACTGGGAATCCCTGAACCCGAGGAGAGATATTTATAT	1907
QY	269	CTCACATCTTAGCTTTTTCAGCAGCAGTGATGGAATTTGATGAAATTTGGTGAGC	328
DB	1906	CCATGTTCTGCTTCTTTTGCAGCAAGCATGGCATAGTAATGAACCTTGGTGAGC	1847
QY	329	GCTTTAGTCAGAGGTGAGGTTCCAGAGGTCGCTGTTTCTATGCGTTTCAAAATCTCA	388
DB	1846	GATTTAGCCAAAGTTCAGATTACAGAGCCGCTGTTTCTATGCGTTTCCAAATTGCCA	1787
QY	389	TCGAGATGTTCACTCAGAGATGACAGTTGCTGATAGACACTTACATCAGAGATCCCA	448
DB	1786	TGAAACACATATCTGAAATGTATAGTCTTATGACACTTACATAAAGATCCCA	1727
QY	449	AGAAAGGGAAATTTTATTAATGCAATTTGAACACCATGCCCTATGTTAAGAAAAAGCAG	508
DB	1726	AGAAAGGGAATTTCTTCAATGCAATTTGAACAGATGCTTGTGTCAGAGAGGCGAG	1667
QY	509	ATTGGGCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	568
DB	1666	ACTGGGCTTGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1607
QY	569	TTGCTGCTGATAGAGAGTTTCTTCTCAGGATCTTTCCTGCTATATTTCTGCTAAGA	628
DB	1606	TTGCTGCTGATAGAGAGTTTCTTCTCAGGATCTTTCCTGCTATATTTCTGCTAAGA	1547
QY	629	AGAGAGGCTTATGCGCAGGACTCCTTTTTCATGAACTCATCAGCAGAGATGAGGAC	688
DB	1546	ACGAGGAGTATGCTGCGCTCACATTTTCTAATGAATTTATAGCAGATGAGGGTT	1487

QY	689	TTCACTGTGACTTTGCTTGCCTGATGTTCCCAATCTTAGTAAATAAGCCTTCAGAGAAA	748
DB	1486	TACACTGTGATTTGCTTGCCTGATGTTCAAAACACCTGGTACACAAACCATCGGAGGAG	1427
QY	749	GGGTACGGGAGATCATTTGATGCTGTCAAAATTTGACGAGGATTTTAAACAGAGCCT	808
DB	1426	GAGTAAGAGAAATAATATCAATGCTGTTCGATAGAAACAGGAGTTCTCTACTGAGGCT	1367
QY	809	TGCCAGTTGGCCTCATTTGGAATGAATTCATTTTGTGATGAAACAGTACATTTGATTTGTAG	868
DB	1366	TGCTGTGAAGCTCATTTGGAATGAATTCATTTTGTGATGAAACAGTACATTTGATTTGTGG	1307
QY	869	CTGACAGATTTACTTGTGAACTTGGATTTCTCAAAAGTTTTCAGCAGCAAAATCTTTTG	928
DB	1306	CAGACAGCTTATGCTGGAATGGTTTTCAGCAAGTTTTCAGAGTAGAGAACCATTTTG	1247
QY	929	ATTTTATGAAAAACATTTCTTTTGAAGAAAAACAAATTTCTTTGAGAAAAAGTTTTCAG	988
DB	1246	ACTTTATGGAATATTTTCACTGGAAGGAAAGACTAACTTCTTTGAGAAAGAGTAGGCG	1187
QY	989	AGTATCAGCGTTTTCAGTTATGTCAGAAACACAGATACGTTCTTACCTTGGATGCG	1048
DB	1186	AGTATCAGAGGATGGGAGTGTCTCAAGTCCAAACAGAGAAATTTCTTTTACCTTGGATGCTG	1127
QY	1049	ATTT 1052	
DB	1126	ACTT 1123	

RESULT 10
ABL65414
ID ABL65414 standard; DNA; 2500 BP.
XX
AC ABL65414;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:3751.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
OS Homo sapiens.
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233333P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.

PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 02-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
DR
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
XX Claim 1; SEQ ID NO 3751; 44pp; English.
PS
XX The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (II) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms
CC tumour
XX
SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;

Query Match 55.1%; Score 580; DB 6; Length 2500;
Best Local Similarity 75.1%; Pred. No. 1.4e-154;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
QY 89 ATGAAGAGCCATCTCTAAGAAAGAGTTCTCGCGGGTTTGTCTATCTTCCAAATCCAGTACC 148
DB 397 AGGATGAGCGGTCTGAGAGAAACCCCGCGGTCTTGTCTATCTTCCCATCGATACC 456
QY 149 CTGATATTGGAAATGTATATAACAGGCACAGCTTCTTCTGGACAGAGAGAGTTG 208
DB 457 ATGATATCTGGCAGATGTATAAGAGGACAGAGCTTCTTCTGGACCGCGAGAGTTG 516
QY 209 ACTTATCAAGATCTCCCTGCTGGAACAAGCTTAAACAGATGAGAGTACTTCTATCT 268
DB 517 ACCTTCCAAGGACATTCAGCATCTGGAAATCCCTGAAACCCGAGGAGATATTTATAT 576
QY 269 CTCACATCTTACCTTTTTTGGACGAGTGAGTGAATTTGTAATGAAATTTGGTGAGC 328

DB 577 CCATATGTTCTGGCTTTCTTTTTCAGCAAGCGATGCGATAGTAATAAATCTGGTGGAGC 636
QY 329 GCTTTAGTCAGGAGGTGCGAGGCTCGCTGTTCTATGGCTTTCAAAATTTCTCA 388
DB 637 GATTTAGCCAAAGTTTCAATTTACAGAGCCGCTGTTCTATGGCTTTCAAAATTTGCCA 696
QY 389 TCGAGAAATGTTCACTCAGAGATGTACAGTTTGTGTAGACACTTATCATCAGAGATCCCA 448
DB 697 TGGAAACATACATTTCTGAAATGTATAGTCTTCTTATTGACACTTACATAAAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTATGCAATGCAATGCAACCATGCCCCATGTTAAGAAAAAGCAG 508
DB 757 AAGAAAGGGAATTTCTCTCAATGCCATGAAACGATGCTTGTGTCAAGAAGAGGCGAG 816
QY 509 ATTGGCGCTTGCAGATGATAGCAGATAGAAATCTACTTTTGGGAAAGAGTGGTGGCT 568
DB 817 ACTGGGCTTTCGCTGATTTGGGACAAAGAGGCTACTATGGTGAACGTGTTGTAGCCT 876
QY 569 TTGCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTGTGCTATATTTGGCTTAAAGA 628
DB 877 TTGCTGCAAGTGAAGGCAATTTCTTTTCCGGTCTTCTTGGCTCGATATTTGGCTCAAGA 936
QY 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGAC 688
DB 937 AACGAGGACTGATGCTGCTGCTCAAAATTTGAGCAGGAGTTTAAACAGAGGCT 996
QY 689 TTCACGTGACTTTGCTTGGCTGATGTTCCAAATCTAGTAATAAAGCCTTCAGAGAAA 748
DB 997 TACACTGTGATTTTGTCTGCTGATGTTCAAAACACCTGGTACAAACCATCGGAGGAGA 1056
QY 749 GGGTCAGGAGGATCATTTGTTGATGCTCTCAAAATTTGAGCAGGAGTTTAAACAGAGGCT 808
DB 1057 GAGTAAGAGAAATAATTAATCAATGCTCTCGGATAGAACAGAGTTCTCTCAGGCT 1116
QY 809 TGCAGTTGGCCTCATTTGGAAATGAATTCATTTTGTATGAGAAACAGTACATTTGAGTTGTAG 868
DB 1117 TGCCTGTGAAGCTCATTTGGGATGAATTTGCACTCTTAATGAAGCAATACATTTGAGTTGTGG 1176
QY 869 CTGACAGTTACTTTGTGGAATCTGATCTCAAAAGTTTTCAGGCAAGAAATCCCTTTG 928
DB 1177 CAGACAGACTTATGCTGGAATCTGGTTTATGAGAGTTTTCAGAGTAGAGAACCCATTTG 1236
QY 929 ATTTTATGAAAAATTTCTTTAGAGAAAAACAAATTTCTTTGAGAAAAACGATTTTCAG 988
DB 1237 ACTTTATGAGATATTTCTGGAAGAAAGACTTAATCTTTTGGAGAGAGATAGGCG 1296
QY 989 AGTATCAGCGTTTGTGCAAGTTATGCGAGAAACCAAGATACGTTTCACTTGGATGAG 1048
DB 1297 AGTATCAGAGGATGGGAGTGTGTCAGTCCACAGAGAAATCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
DB 1357 ACTT 1360
RESULT 11
ABL66517
ID ABL66517 standard; DNA; 2500 BP.
XX
XX ABL66517;
XX AC
XX AC
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:4854.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilms's tumour; adenocarcinoma;
XX gene; ds.
OS Homo sapiens.
XX
PN WO200194629-A2.

XX 13-DEC-2001.
 XX PD
 XX PF
 XX PR 30-MAY-2001; 2001WO-US010838.
 XX PR 05-JUN-2000; 2000US-0209473P.
 XX PR 05-JUN-2000; 2000US-0209531P.
 XX PR 18-SEP-2000; 2000US-0233133P.
 XX PR 18-SEP-2000; 2000US-0233617P.
 XX PR 20-SEP-2000; 2000US-0234009P.
 XX PR 20-SEP-2000; 2000US-0234034P.
 XX PR 20-SEP-2000; 2000US-0234052P.
 XX PR 22-SEP-2000; 2000US-0234509P.
 XX PR 22-SEP-2000; 2000US-0234567P.
 XX PR 25-SEP-2000; 2000US-0234923P.
 XX PR 25-SEP-2000; 2000US-0234924P.
 XX PR 25-SEP-2000; 2000US-0235077P.
 XX PR 25-SEP-2000; 2000US-0235082P.
 XX PR 25-SEP-2000; 2000US-0235134P.
 XX PR 25-SEP-2000; 2000US-0235280P.
 XX PR 26-SEP-2000; 2000US-0235637P.
 XX PR 26-SEP-2000; 2000US-0235638P.
 XX PR 27-SEP-2000; 2000US-0235711P.
 XX PR 27-SEP-2000; 2000US-0235720P.
 XX PR 27-SEP-2000; 2000US-0235840P.
 XX PR 27-SEP-2000; 2000US-0235863P.
 XX PR 28-SEP-2000; 2000US-0236028P.
 XX PR 28-SEP-2000; 2000US-0236032P.
 XX PR 28-SEP-2000; 2000US-0236033P.
 XX PR 28-SEP-2000; 2000US-0236034P.
 XX PR 28-SEP-2000; 2000US-0236109P.
 XX PR 28-SEP-2000; 2000US-0236111P.
 XX PR 29-SEP-2000; 2000US-0236842P.
 XX PR 29-SEP-2000; 2000US-0236891P.
 XX PR 02-OCT-2000; 2000US-0237172P.
 XX PR 02-OCT-2000; 2000US-0237173P.
 XX PR 02-OCT-2000; 2000US-0237278P.
 XX PR 02-OCT-2000; 2000US-0237294P.
 XX PR 02-OCT-2000; 2000US-0237295P.
 XX PR 02-OCT-2000; 2000US-0237316P.
 XX PR 03-OCT-2000; 2000US-0237425P.
 XX PR 03-OCT-2000; 2000US-0237598P.
 XX PR 03-OCT-2000; 2000US-0237604P.
 XX PR 03-OCT-2000; 2000US-0237606P.
 XX PR 03-OCT-2000; 2000US-0237608P.
 XX PR 01-NOV-2000; 2000US-0244867P.
 XX PR 01-NOV-2000; 2000US-0245084P.
 XX PA (AVAL-) AVALON PHARM.
 XX PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 XX PI Soppet DR, Weaver Z;
 XX DR WPI; 2002-188264/24.
 XX PT Screening for anti-neoplastic agent involves exposing cells to a chemical
 XX PT agent to be tested for anti-neoplastic activity, and determining a change
 XX PT in expression of a gene of a signature gene set.
 XX PS Claim 1; SEQ ID NO 4854; 44pp; English.
 XX CC The present invention describes a method (M1) for screening for an anti-
 XX CC neoplastic agent. The method involves exposing cells to a chemical agent
 XX CC to be tested for anti-neoplastic activity, determining a change in
 XX CC expression of at least one gene (I) of a signature gene set, where (I)
 XX CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 XX CC to ABL70110), or is at least 95% identical to (S), where a change in
 XX CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 XX CC activity and can be used in gene therapy. M1 can be used for screening an
 XX CC anti-neoplastic agent, and can be used for producing a product which is
 XX CC the data collected with respect to the anti-neoplastic agent as a result
 XX CC of M1, and the data is sufficient to convey the chemical structure and/or
 XX CC properties of the agent. M1 can be used in the treatment of cancer such

CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
 CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
 CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
 CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
 CC tumour
 XX
 SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;
 Query Match 55.1%; Score 580; DB 6; Length 2500;
 Best Local Similarity 75.1%; Pred. No. 1.4e-154;
 Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
 QY 89 ATCAAGAGCCACTCTAAGAAAGAGTTCTGCCCGTTTGTCTATCTTTCCATCCAGTACC 148
 DB |||||
 QY 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCCGCTTGTTCATCTTCCCATCGAGTACC 456
 DB |||||
 QY 149 CTGATATTTGGAAAAATGATATAAAGCAGCAGGCTTCTTCTGGACAGCAGAGAGGTTG 208
 DB |||||
 QY 457 ATGATATCTGGCAGATGATATAAGAGGAGGAGGCTTCTTTGGACCGCGAGGAGTTG 516
 DB |||||
 QY 209 ACTTATCAAGAGATCTCCCTCACTGGAAACAGCTTAAGCAGATGAGAGTACTTCACT 268
 DB |||||
 QY 517 ACCTCTCCAAGGACATTCAGACCTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
 DB |||||
 QY 269 CTCACATCTTAGCCTTTTTCAGCCAGTATGAAATGTAAATCAAAATTTGGTGGAGC 328
 DB |||||
 QY 577 CCATGTTCTGGCTTCTTTCAGCAAGCAGATGCAATGATTAATGAATCTTGGTGGAGC 636
 DB |||||
 QY 329 GCTTTAGTCAGGAGGTCAGGTTTCAGAGGCTCGCTTCTTCTATGGCTTCAAAATTTCA 388
 DB |||||
 QY 637 GATTTAGCCAAGAAGTTTCAGATTACAGAGCCGCTTCTTCTATGGCTTCAAAATTTGCA 696
 DB |||||
 QY 389 TCGAGAAATGTTCACTCAGAGATGACAGTTTCTGTATAGACACTTACATCAGAGATCCCA 448
 DB |||||
 QY 697 TGGAAAAACATACATCTCTGAAATGTATAGTCTTCTTATTGACACTTACATAAAGATCCCA 756
 DB |||||
 QY 449 AGAAAAGGGAATTTTATTAATGCAATGAAACCATGCCCCTATGTTAAAGAAAAAGCAG 508
 DB |||||
 QY 757 AGAAAAGGGAATTTCTTCAATGCCATTAAGAGATGCTTGTCTAAGAAAGGAGCAG 816
 DB |||||
 QY 509 ATTGGGCTTTCGATGATAGCAGATAGAAAATCTACTTTTGGGAAAGAGTGGTGGCT 568
 DB |||||
 QY 817 ACTGGGCTTTCGCTGGATGGGAGCAAGAGGCTACTATGGTGAACGTTGTAGCT 876
 DB |||||
 QY 569 TTGCTGCTGTAGAGGAGTTTCTTCTCAGAACTTTTGTGCTGCTATTTCTGGCTAAAGA 628
 DB |||||
 QY 877 TTGCTGAGTGAAGGCAATTTCTTTTCCGCTTCTTTGCGTCTGATATTTCTGCTCAAGA 936
 DB |||||
 QY 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCATGAACTCATCAGCAGATGAGGAC 688
 DB |||||
 QY 937 AAGCAGGAGTATGCTGGCTTCACTATTTTCTAATGAACCTTATAGCAGATGAGGCT 996
 DB |||||
 QY 689 TTCACCTGTGATTTTGTCTGCTGATGTTCCAATCTTCTAGTAAATTAAGCTTTCAGAGAAA 748
 DB |||||
 QY 997 TACACTGTGATTTTGTCTGCTGATGTTTCAAAACACCTGGTACACAAACCATCGAGGAGA 1056
 DB |||||
 QY 749 GGGTCAGGAGATCATTTGTGATGCTGTCAAAATTTGAGCAGGAGTTTAAACAAGAGCT 808
 DB |||||
 QY 1057 GAGTAAGAGAAATAATATCAATGCTGTTCGGATAGAAACAGGAGTTCTCTACTGAGGCT 1116
 DB |||||
 QY 809 TGGCAGTTGGCTCATTTGGAATGAAATGCAATTTTGTATGAACAGTACATTGAGTTTGTAG 868
 DB |||||
 QY 1117 TGCCTGTGAAGCTCATTTGGGATGAATTTGCACTCTAATGAAGCAATACATTGAGTTTGTGG 1176
 DB |||||
 QY 869 CTGACAGATTTACTTGTGGAACTTGGATTCTCAAGGTTTTCAGGCGAGAAATCTTTTG 928
 DB |||||
 QY 1177 CAGACAGACTATGCTGGAACTCGGTTTTCAGAGAGTTTTCAGAGTGAAGAACCATTTG 1236
 DB |||||
 QY 929 ATTTTATGAAAAACATTTCTTTTGAAGGAAAAACAAATTTCTTTGAGAAAACGAGTTTCAG 988
 DB |||||
 QY 1237 ACTTTATGAGAAATATTTCACTGGAAGAAAGACTTAATCTTTTGAAGAGAGTAGGCG 1296
 DB |||||
 QY 989 AGTATCAGCGTTTTCAGAGTTATGGCAGAAACACAGATAACGTTCTTCACTTGGATGCGAG 1048
 DB |||||

Db 1297 ACTATCAGAGGATGGAGTGATGTCAGGTCACAGAGAAATCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 12

ABL65859
ID ABL65859 standard; DNA; 2500 BP.

XX ABL65859;
XX 15-MAY-2002 (first entry)
XX Lung cancer related gene sequence SEQ ID NO:4196.
XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilms tumour; adenocarcinoma;
KW gene; db.

OS Homo sapiens.

XX WO200194629-A2.

XX 13-DEC-2001.

XX 30-MAY-2001; 2001WO-US010838.

XX 05-JUN-2000; 2000US-0209473P.

XX 05-JUN-2000; 2000US-0209531P.

XX 18-SEP-2000; 2000US-0233133P.

XX 20-SEP-2000; 2000US-0233617P.

XX 20-SEP-2000; 2000US-0234009P.

XX 20-SEP-2000; 2000US-0234034P.

XX 20-SEP-2000; 2000US-0234052P.

XX 22-SEP-2000; 2000US-0234509P.

XX 22-SEP-2000; 2000US-0234567P.

XX 25-SEP-2000; 2000US-0234923P.

XX 25-SEP-2000; 2000US-0234924P.

XX 25-SEP-2000; 2000US-0235077P.

XX 25-SEP-2000; 2000US-0235082P.

XX 25-SEP-2000; 2000US-0235134P.

XX 25-SEP-2000; 2000US-0235280P.

XX 26-SEP-2000; 2000US-0235637P.

XX 26-SEP-2000; 2000US-0235711P.

XX 27-SEP-2000; 2000US-0235720P.

XX 27-SEP-2000; 2000US-0235840P.

XX 27-SEP-2000; 2000US-0235863P.

XX 28-SEP-2000; 2000US-0236028P.

XX 28-SEP-2000; 2000US-0236032P.

XX 28-SEP-2000; 2000US-0236033P.

XX 28-SEP-2000; 2000US-0236034P.

XX 28-SEP-2000; 2000US-0236109P.

XX 29-SEP-2000; 2000US-0236111P.

XX 29-SEP-2000; 2000US-0236842P.

XX 29-SEP-2000; 2000US-0236891P.

XX 02-OCT-2000; 2000US-0237172P.

XX 02-OCT-2000; 2000US-0237173P.

XX 02-OCT-2000; 2000US-0237278P.

XX 02-OCT-2000; 2000US-0237294P.

XX 02-OCT-2000; 2000US-0237295P.

XX 02-OCT-2000; 2000US-0237316P.

XX 03-OCT-2000; 2000US-0237425P.

XX 03-OCT-2000; 2000US-0237598P.

XX 03-OCT-2000; 2000US-0237604P.

XX 03-OCT-2000; 2000US-0237606P.

XX 01-NOV-2000; 2000US-0237608P.

XX 01-NOV-2000; 2000US-0244867P.

XX 01-NOV-2000; 2000US-0245084P.

XX (AVAL-) AVALON PHARM.
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
XX Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.
XX Screening for anti-neoplastic agent involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, and determining a change
XX in expression of a gene of a signature gene set.
XX Claim 1; SEQ ID NO 4196; 44pp; English.
XX The present invention describes a method (M1) for screening for an anti-
XX neoplastic agent. The method involves exposing cells to a chemical agent
XX to be tested for anti-neoplastic activity, determining a change in
XX expression of at least one gene (I) of a signature gene set, where (I)
XX comprises a sequence (S) selected from 8447 sequences (given in ABL61664
XX to ABL70110), or is at least 95% identical to (S), where a change in
XX expression is indicative of anti-neoplastic activity. (I) has cytostatic
XX activity and can be used in gene therapy. M1 can be used for screening an
XX anti-neoplastic agent, and can be used for producing a product which is
XX the data collected with respect to the anti-neoplastic agent as a result
XX of M1, and the data is sufficient to convey the chemical structure and/or
XX properties of the agent. M1 can be used in the treatment of cancer such
XX as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
XX prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
XX cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
XX cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms
XX tumour
XX Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;
Query Match 55.1%; Score 580; DB 6; Length 2500;
Best Local Similarity 75.1%; Pred. No. 1.4e-154;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
QY 89 ATGAAGAGCCACTCTTAAGAAAGAGTTCTGCCGGTTTGTCATCTTTCCATCCAGTACC 148
Db 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCCGCTTGTGTCATCTTCCCATCGAGTACC 456
QY 149 CTGATATTTGGAATAATGTATAAAGCAGGCACAGGCTTCTCTGACACAGAGAGGTTG 208
Db 457 ATGATATCTGCGAGATGTATAAAGGCGAGGCTTCTTTTGGACCCCGAGGAGTTG 516
QY 209 ACTTATCAAGAGTCTCCCTCACTGGAAACAAAGCTTAAAGCAGATGAGAAGTACTTCATCT 268
Db 517 ACCTCTCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
QY 269 CTCACATCTTAGCTTTTTCGAGCCAGTGATGGAATGTAAATGAAATTTGGTGGAGC 328
Db 577 CCATGTTCTGGCTTTCTTTGCGAGCAAGCGATGCGATGAAATGAAATCTTGGTGGAGC 636
QY 329 GCTTTAGTCAGGAGTGCAGGCTCCAGAGGCTCGCTGTTTCTATGGCTTCAAAATTTCTCA 388
Db 637 GATTTAGCCAAAGAGTTTCAGATTACAGAGCCCGCTGTTTCTATGGCTTCAAAATTTGCCA 696
QY 389 TCGAGATGTTCACTCAGAGATGTACAGTTTGTGCTAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAACATACATTCTGAAATGTATAGTCTTCTTATTGACACTTACATAAAGATCCCA 756
QY 449 AGAAAAGGGAATTTTATTAAATGCAATGAAACCATGCCCTATGTTTAAAGAAAAAGAG 508
Db 757 AAGAAAGGGAATTTCTCTCAATGCCATTGAAACGATGCCCTTGTGTCAAGAAAGAGGAG 816
QY 509 ATTGGGCTTGGCATGATAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCT 568
Db 817 ACTGGGCTTGGCGCTGGATTGGGGACAAAGAGGCTCTCTATGGTGAACGTGTTAGCT 876
QY 569 TTGCTGCTGTAAGAGGAGTTTCTTCTCAGGATCTTTTGTGCTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGCAAGTGGAAAGGCAATTTCTTTTCCGGTCTTTTTCGTCGATATTTCTGGCTCA 936

QY 569 TTGCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTGTCTGCTATATCTTGCTAAAGA 628
Db 877 TTGCTGCTGTAGAAGGAGTTTCTTCTTCCGTTCTTTTGGTCTGATATCTTGCTCAAGA 936
QY 629 AGAGAGGCTTATGCCAGGACTCATTCTTCCATGAACTCATCAGCAGAGATGAAGAC 688
Db 937 AACGAGGACTGATGCTGCGCTCACATTTTCTTAATGAACCTTATAGCAGAGATGAGGGTT 996
QY 689 TTCACTGTCACTTTGCTGCTGATGCTTCCAACTACTAGTAATAAGCCTTTCAGAGAA 748
Db 997 TACACTGTGATTTGCTGCTGCTGATGCTTCAACACTGTTACAAACCATCGGAGAGA 1056
QY 749 GGGTCAGGAGATCAATTTGTGATGCTGTCAAAATTTGAGCAGGAGTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATAATTATCAATGCTGTTGCGATAGAACAGGAGTTTCTCACTGAGGCT 1116
QY 809 TGCAGTTGGCTCATTTGAATGAATGCAATTTTGTATGAAACAGTACATGAGTTGTAG 868
Db 1117 TGCCTGTGAAGCTCATTTGGGATGAATTGCACTCTAATGAAGCAATACATTTGAGTTGTGG 1176
QY 869 CTGACAGATTACTGTGGAACCTTGATTTCTCAAGGCTTTTTCAGGCAGAAATCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACCTGGGTTTATGCAAGGTTTTCAGATAGAGAACCCATTG 1236
QY 929 ATTTATGAAACCAATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTTCAG 988
Db 1237 ACTTTATGAGAAATTTTCACTGGAAGGAAAGCTAACTTTCTTTGAGAGAGATGAGCG 1296
QY 989 AGTATCAGCGTTTGTGAGTTATGGCAGAGAAACACAGATAACGTTCTTCACTTGGAGCAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGATGTCAAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360
RESULT 14
ADI32044
ID ADI32044 standard; cDNA; 2500 BP.
XX
AC ADI32044;
DT 17-JUN-2004 (first entry)
XX
DE Human cDNA #1370.
XX
KW Human; gene; ss; immunological response; immunopathological condition;
KW Crohn's disease; asthma; ulcerative colitis; hypereosinophilia;
KW irritable bowel syndrome; osteoarthritis; rheumatoid arthritis;
KW acute monocytic leukaemia; antinflammatory; antidiabetic; antiulcer;
KW osteopathic; antiarthritic; antirheumatic; cytostatic.
XX
OS Homo sapiens.
XX
PN US6607879-B1.
XX
PD 19-AUG-2003.
XX
PF 09-FEB-1998; 98US-00023655.
XX
PR 09-FEB-1998; 98US-00023655.
XX
PA (INCY-) INCYTE CORP.
XX
PI Cocks BG, Stuart SG, Seilhamer JU;
XX
DR WPI; 2003-895307/82.
XX
PT A composition comprising a plurality of cDNAs, useful for detecting
PT altered expression of genes in an immunological response or for
PT diagnosing and treating an immunopathology, e.g. Crohn's disease, asthma
PT or osteoarthritis.

XX
PS
XX
CC The invention relates to a composition comprising a plurality of cDNAs
CC for detecting the altered expression of genes in an immunological
CC response. The invention also relates to a method of diagnosing or
CC monitoring the treatment of an immunopathological condition in a sample,
CC comprising obtaining nucleic acids from a sample, contacting the nucleic
CC acids of the sample with an array comprising the plurality of cDNAs under
CC conditions to form one or more hybridisation complexes, detecting the
CC hybridisation complexes and comparing the levels of the detected
CC hybridisation complexes with the level of hybridisation complexes
CC detected in a non-diseased sample, where an altered level of the detected
CC hybridisation complexes correlates with the presence of an
CC immunopathological condition. Also disclosed are an expression profile
CC comprising a microarray and a plurality of detectable complexes and a
CC method for identifying a plurality of polynucleotide probes. The cDNAs
CC are useful as hybridisable array elements in a microarray for monitoring
CC the expression of target polynucleotides. The microarray can be used in
CC the diagnosis of an immunopathology, such as Crohn's disease, asthma,
CC ulcerative colitis, hypereosinophilia, irritable bowel syndrome,
CC osteoarthritis, rheumatoid arthritis or acute monocytic leukaemia, and in
CC identifying agents for the treatment of the diseases. The microarray may
CC also be used in drug discovery and development, toxicological and
CC carcinogenicity studies, forensics or pharmacogenomics. The composition
CC may also be used in purification of a subpopulation of mRNAs, cDNAs or
CC genomic fragments. This sequence represents a human cDNA of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification but was obtained in electronic format directly
CC from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;
Query Match 55.1%; Score 580; DB 11; Length 2500;
Best Local Similarity 75.1%; Pred. No. 1.4e-154;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
QY 89 ATGAAGAGCCACTCTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTTCCAAATCCAGTACC 148
Db 397 AGGATGAGCGCTGTGTGAGAGAAACCCCGCGCTTTGTGTCATCTTCCCATCCAGTACC 456
QY 149 CTGATATTTGAAAAATGTATAAAACAGGCACAGGCTTCTTTCTGGACAGCAGAGAGTTG 208
Db 457 ATGATATCTGGCAGATGTATAAGAGGCGCAGAGGCTTCTTTTGGACCGCCGAGAGTTG 516
QY 209 ACTTATCAAGAGATCTCCCTCACTGGAACAAAGCTTAAAGCAGATGAGAGTACTTCATCT 268
Db 517 ACCTCTCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
QY 269 CTCACATCTTAGCCTTTTTCAGCCAGTCATGGAATGTAAATGAAATTTGGTGGAGC 328
Db 577 CCCATGTTCTGGCTTCTTTTGCAGCAAGCCATGATGATAGTAATAAGAACTTTGGTGGAGC 636
QY 329 GCTTTAGTCAGGAGGTGCGAGTTCAGAGGCTCGCTTCTTCTATGGCTTTCAAAATTTCTCA 388
Db 637 GATTAGCCAAGAGTTTCAGATTACAGAACCCGCTCTTCTATGGCTTCCAAATGGCA 696
QY 389 TCAGAAATGTTCTACAGAGATGTACAGTTTGTGTGATAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAAAACATACATTTCTGAAATGTATAGTCTTCTTATTGACACTTACATAAAGATCCCA 756
QY 449 AGAAAGGGAAATTTTATTTAATGCAATTCGAACCATGCCCCCTATGTTAAGAAAAAGCAG 508
Db 757 AAGAAAGGGAAATTTCTCTTCAATGCCATTTGAAACGATGCGCTTGTGCAAGAAAGGCGAG 816
QY 509 ATTGGGCTTGGCATGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCT 568
Db 817 ACTGGGCTTGGCTGATTTGGGACAAAGAGGCTACTATGGTGAACGCTTTGTAGGCT 876
QY 569 TTGCTGCTGTAGAAGGAGTTTCTTCTCAGAGATCTTTTGTGCTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGAGGAGGAGGAGTCTTTTCTTCCGCTTCTTTTGGCTCGATATTTCTGGCTCAAGA 936

Db 1357 ACTT 1360

Search completed: October 30, 2005, 00:02:46
Job time : 442.5 secs

Result No.	Query %			DB	ID	Description
	Score	Match	Length			
1	1051.4	99.8	3292	3	CR617553	CR617553 full-length
2	1051.4	99.8	4748	3	HSW802033	ALI37348 Homo sapi
3	1039.4	98.7	4650	3	BC042948	BC042948 Homo sapi
4	1009.9	95.8	3501	3	CR627376	CR627376 Homo sapi
5	943.4	89.6	1056	9	AV398973	AV398973 Homo sapi
6	928.8	88.1	1083	1	AL547501	AL547501
7	828.2	78.7	1008	9	AV398974	AV398974 Pan trogl
8	791.6	75.2	879	5	BQ441857	BQ441857 AGENCOURT
9	778.6	73.9	1043	9	AV398975	AV398975 Mus muscu
10	731.2	69.4	892	7	CF163214	CF163214 952695 MA
11	726.6	69.0	914	7	CF595079	CF595079 AGENCOURT
12	710.6	67.5	870	7	CO648380	CO648380 ILLUMINIGEN
13	682.8	64.8	804	6	DC656661	DC656661 AGENCOURT
14	653.8	62.1	1061	4	BM801298	BM801298 AGENCOURT
15	653.2	62.0	896	5	BU170979	BU170979 AGENCOURT
16	614.8	58.4	685	5	EX951374	EX951374 DKP2p781E
17	578.4	54.9	1571	3	CR603461	CR603461 full-length
18	578.4	54.9	1573	3	CR625489	CR625489 full-length
19	578.4	54.9	1582	3	CR608076	CR608076 full-length
20	578.4	54.9	1588	3	CR602054	CR602054 full-length
21	578.4	54.9	1592	3	CR621427	CR621427 full-length
22	578.4	54.9	1600	3	CR604378	CR604378 full-length
23	578.4	54.9	1605	3	CR6590359	CR6590359 full-length
24	578.4	54.9	1612	3	CR614990	CR614990 full-length

QY	61	GACACCAACGAAAGTGAAATGAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTCTTCGC	120
Db	87	GACACCAACGAAAGTGAAATGAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTCTTCGC	146
QY	121	CGGTTTGTCATCTTTCCAAATCCAGTACCCCTGATATTTGGAAATGTATATAACAGGCACAG	180
Db	147	CGGTTTGTCATCTTTCCAAATCCAGTACCCCTGATATTTGGAAATGTATATAACAGGCACAG	206
QY	181	GCTTCCTCTGACACAGAGAGGTGACTTATCAAGATCTCCCTCACTGGAAACAG	240
Db	207	GCTTCCTCTGACACAGAGAGGTGACTTATCAAGATCTCCCTCACTGGAAACAG	266
QY	241	CTTAAAGCAGATGAGAAGTACTCTCTCTACATCTTAGCCCTTTTTCAGCCAGTGAT	300
Db	267	CTTAAAGCAGATGAGAAGTACTCTCTCTACATCTTAGCCCTTTTTCAGCCAGTGAT	326
QY	301	GGAAATGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCAGAGGCT	360
Db	327	GGAAATGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCAGAGGCT	386
QY	361	CGTGTGTTCTATGGCTTTCAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG	420
Db	387	CGTGTGTTCTATGGCTTTCAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG	446
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATCAATTTGAA	480
Db	447	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATCAATTTGAA	506
QY	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGATAGATAGAGAA	540
Db	507	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGATAGATAGAGAA	566
QY	541	TCATCTTTTGGGAAAGAGTGGGCTTTGCTGTGTAGAGAGTTCCTCTCAGGA	600
Db	567	TCATCTTTTGGGAAAGAGTGGGCTTTGCTGTGTAGAGAGTTCCTCTCAGGA	626
QY	601	TCCTTTGCTGCTATATCTGGCTTAAAGAGAGAGGTCTTATGCCAGAGCTCACTTTTCC	660
Db	627	TCCTTTGCTGCTATATCTGGCTTAAAGAGAGAGGTCTTATGCCAGAGCTCACTTTTCC	686
QY	661	AATGAACCTCATCAGCAGATGAAGACTTCCTGCTGCTTTGCTTGCCTGATGTTCCAA	720
Db	687	AATGAACCTCATCAGCAGATGAAGACTTCCTGCTGCTTTGCTTGCCTGATGTTCCAA	746
QY	721	TACTTAGTAAATGAAGCTTTCAAGAAAGAGGTGAGGATCATGCTGCTGTCCTCA	780
Db	747	TACTTAGTAAATGAAGCTTTCAAGAAAGAGGTGAGGATCATGCTGCTGTCCTCA	806
QY	781	ATTGAGCAGAGTCTTTTAAACAGAGCTTGCAGTTGGCTCATTTGGAATGAATTCAT	840
Db	807	ATTGAGCAGAGTCTTTTAAACAGAGCTTGCAGTTGGCTCATTTGGAATGAATTCAT	866
QY	841	TTGATGAACACAGTACATGAGTTGTAGCTGACAGATTAATCTGTGAACTTGGATCTCA	900
Db	867	TTGATGAACACAGTACATGAGTTGTAGCTGACAGATTAATCTGTGAACTTGGATCTCA	926
QY	901	AAGGTTTTTCAGGAGAAATCTTTTGAATTTTATGGAACAACTTTCTTTAGAGGAAA	960
Db	927	AAGGTTTTTCAGGAGAAATCTTTTGAATTTTATGGAACAACTTTCTTTAGAGGAAA	986
QY	961	ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAAC	1020
Db	987	ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAAC	1046
QY	1021	ACAGATAAGCTCTTCACTTGGATGAGATTTT	1053
Db	1047	ACAGATAAGCTCTTCACTTGGATGAGATTTT	1079
RESULT 2			
HSMB02033			
LOCUS	HSMB02033	4748 bp	mRNA linear HTC 22-SEP-2004

DEFINITION	Homo sapiens mRNA; cDNA DKFp761E1312 (from clone DKFp761E1312).
ACCESSION	AL137348
VERSION	AL137348.1
KEYWORDS	HTC.
SOURCE	GI:6807859
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 4748)
CONSTRM	Ansorge,W., Krieger,S., Regiert,T., Rittmueller,C., Schwager,B., Wewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
TITLE	The German cDNA Consortium
JOURNAL	Direct Submission
COMMENT	Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by EMBL (European Molecular Biology Laboratories, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFp761E1312) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFp761E1312 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.
FEATURES	Location/Qualifiers
source	1..4748
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="RZPD:DKFp761E1312"
	/db_xref="taxon:9606"
	/clone="DKFp761E1312"
	/tissue_type="amygdala"
	/clone_lib="761 (synonym: hamy2). Vector pSport1; host DH10B; sites NotI + SalI"
	/dev_stage="adult"
	/notes="ribonucleotide reductase"
gene	1..4748
CDS	/genes="DKFp761E1312"
	47..1102
	/genes="DKFp761E1312"
	/codon_start=1
	/product="hypothetical protein"
	/protein_id="CAB70703.2"
	/db_xref="GI:52545612"
	/db_xref="GOA:Q9NTD8"
	/translation="MGDPPEAAGLDQDRSSSDTNESEIKSNEEPLLRKSSRRFVI FPIQYDDIWMYKQQAQSFMTAEEDVLSKDLPHWKNKADDEKIFISHILAFASDGI VNEILVERFSQEVQVPEARCFYQILINHEMYSLLIDTYIRDPKREFLPNAIE TMPYKKADWLRIADRKSTFCERVFAAEEGVFSGFAAI FWLKKRGLMPGLT FSNELISRDGLHCDFAELMFQVNVNKSSEERVREIIVDAVKIQEPLTEALPVLIG MNCILMKQVTEFVADRLVLELGFSGVQAENPDPFMENISLEKTNFEKRVSEYQRF AVMAETDNTVFLDADF"
ORIGIN	
Query Match	99.8%; Score 1051.4; DB 3; Length 4748;
Best Local Similarity	99.9%; Pred. No. 1.6e-268;
Matches 1052; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
QY	1 ATGGCGCAGCCCGAAAGCGCGGCGCGGCGTGGATCAGGATGAGAGATCATCTTCA 60
Db	47 ATGGCGCAGCCCGAAAGCGCGGCGCGGCGTGGATCAGGATGAGAGATCATCTTCA 106
QY	61 GACACCAACGAAAGTGAATATAAGTCAATCAAGAGCCACTCTCTAAGAAAGTCTTCGC 120
Db	107 GACACCAACGAAAGTGAATATAAGTCAATCAAGAGCCACTCTCTAAGAAAGTCTTCGC 166
QY	121 CGGTTTGTCTCATCTTTCCAAATCCAGTACCCCTGATATTTGGAAATGTATATAACAGGCACAG 180
Db	167 CGGTTTGTCTCATCTTTCCAAATCCAGTACCCCTGATATTTGGAAATGTATATAACAGGCACAG 226

```
QY 181 GCTTCCTCTTGGACAGACAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
Db 227 GCTTCCTCTTGGACAGACAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 286
QY 241 CTTAAGCAGATGAGAGTACTTCTCTCTCACTCTTACCTTTTGGAGCAGTGAT 300
Db 287 CTTAAGCAGATGAGAGTACTTCTCTCTCACTCTTACCTTTTGGAGCAGTGAT 346
QY 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db 347 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 406
QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 407 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 466
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATGAA 480
Db 467 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATGAA 526
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTGGATGATACGAGTAGAAA 540
Db 527 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTGGATGATACGAGTAGAAA 586
QY 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGTAGAGAGGTTTCTTCTCAGGA 600
Db 587 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGTAGAGAGGTTTCTTCTCAGGA 646
QY 601 TCTTTTGTCTATATCTTGGCTTAAAGAGAGAGGTCTTATGCCAGACTCACTTTTTC 660
Db 647 TCTTTTGTCTATATCTTGGCTTAAAGAGAGAGGTCTTATGCCAGACTCACTTTTTC 706
QY 661 AATGAACATCAGCAGATGAGACATCTGCTGCTGTAGAGAGGTTTCTTCTCAGGA 720
Db 707 AATGAACATCAGCAGATGAGAGACTTCACTGTGACTTTGCTTGGCTGTGATGTC 766
QY 721 TACTTAGTAAATAAGCTTTCAGAGAAAGGTCAGGAGATCATCTGTCATGCTGTGCA 780
Db 767 TACTTAGTAAATAAGCTTTCAGAGAAAGGTCAGGAGATCATCTGTCATGCTGTGCA 826
QY 781 ATTGAGCAGAGTCTTTTAAACAGAGCTTGGCCTCATTTGGAATGAATTCATTT 840
Db 827 ATTGAGCAGAGTCTTTTAAACAGAGCTTGGCCTCATTTGGAATGAATTCATTT 886
QY 841 TTGATGAAACAGTACATTTAGTTTGTAGCTGACAGATTAATCTGTGAACTTGGATTC 900
Db 887 TTGATGAAACAGTACATTTAGTTTGTAGCTGACAGATTAATCTGTGAACTTGGATTC 946
QY 901 AAGGTTTTTCAGGCAGAAATCTTTTGAATTTATGGAACAATTTCTTTAGAGGAAA 960
Db 947 AAGGTTTTTCAGGCAGAAATCTTTTGAATTTATGGAACAATTTCTTTAGAGGAAA 1006
QY 961 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTGTGAGTATGCGAGAAC 1020
Db 1007 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTGTGAGTATGCGAGAAC 1066
QY 1021 ACAGATAACGTTTCACTTGGATGAGATTTT 1053
Db 1067 ACAGATAACGTTTCACTTGGATGAGATTTT 1099
```

BC042948
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Homo sapiens cdna clone IMAGE:4798175, containing frame-shift errors
BC042948
BC042948.2
GI:34194000
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS

Mammalia; Rutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 4650)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, I., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buotow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Sapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullan, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahy, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalil, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932

TITLE

JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 4650)
Director MGC Project.
Direct Submission
Submitted (09-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NTH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 25, 2003 this sequence version replaced gi:27695575.
Contact: MGC help desk
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabh,
Parvaneh Saeedi, JR Santos, Angeli Schnerch, Ursula Skalska,
Duane Smalil, Jeff Stott, Miranda Tsai, George Yang, Jacque
Schein, Asim Siddiqui, Rob Holt, Marco Marra.

REMARK
COMMENT

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 70 Row: h Column: 19
This clone has the following problem: frame shifted.

FEATURES

source

1. 4650
/organism="Homo sapiens"
/mol type="mRNA"
/db xref="taxon:9606"
/clone="IMAGE:4798175"
/tissue type="Brain, hypothalamus"
/clone lib="NIH_MGC_96"
/lab host="DH10B"
/note="Vector: pBluescript"

ORIGIN

Query Match 98.7%; Score 1039.4; DB 3; Length 4650;
Best Local Similarity 99.8%; Pred. No. 2.5e-265;
Matches 1051; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY

1 ATGGCGGACCCGAAAGCGCGGCGGCTGATCAGATGAGATGAGATCATCTTCA 60
|||||

Db 72 ATGGCGCAGCCCGAAGGCGGAGCGCGCGCTGGATCAGGATGAGAGATCATCTTCA 131
Qy 61 GACACCAAGCAAGTGAATAAGTCAATGAAGACCACTCTCTAAGAAGAGTTCCTGC 120
Db 132 GACACCAAGCAAGTGAATAAGTCAATGAAGAG-CACTCTAAGAAGAGTTCCTGC 190
Qy 121 CGGTTTGTCACTTTCCAAATCCAGTACCCTGTATTTGGAAATGTATAAACAGGCACAG 180
Db 191 CGGTTTGTCACTTTCCAAATCCAGTACCCTGTATTTGGAAATGTATAAACAGGCACAG 250
Qy 181 GCTTCCTTCTGACAGCAGAGAGGTGTGATTTCAAGAGTCTCCCTCACTGGAAACAG 240
Db 251 GCTTCCTTCTGACAGCAGAGAGGTGTGATTTCAAGAGTCTCCCTCACTGGAAACAG 310
Qy 241 CTTAAGCAGATGAGAAGTACTTCTCTCTACATCTTAGCCCTTTTTCGACCCAGTGAT 300
Db 311 CTTAAGCAGATGAGAAGTACTTCTCTCTACATCTTAGCCCTTTTTCGACCCAGTGAT 370
Qy 301 GGAATTTGAAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
Db 371 GGAATTTGAAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 430
Qy 361 CGCTGTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCATCATGAGATGTACAGTTTG 420
Db 431 CGCTGTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCATCATGAGATGTACAGTTTG 490
Qy 421 CTGATAGACATTCATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
Db 491 CTGATAGACATTCATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 550
Qy 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTCGATGATAGCAGATAGAAAA 540
Db 551 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTCGATGATAGCAGATAGAAAA 610
Qy 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAGAAGAGTTCCTTCTCAGGA 600
Db 611 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAGAAGAGTTCCTTCTCAGGA 670
Qy 601 TCTTTTGTGCTATATCTGGCTTAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Db 671 TCTTTTGTGCTATATCTGGCTTAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 730
Qy 661 AATGAACCTCATCAGCAGAGATGAAGACTTCTACTGTGACTTTGCTTGGCTGATGTCCAA 720
Db 731 AATGAACCTCATCAGCAGAGATGAAGACTTCTACTGTGACTTTGCTTGGCTGATGTCCAA 790
Qy 721 TACTTTAGTAAATPAGCTTTCAGAAAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAA 780
Db 791 TACTTTAGTAAATPAGCTTTCAGAAAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAA 850
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTGGCTCATTTGGAATGAATTGCATT 840
Db 851 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTGGCTCATTTGGAATGAATTGCATT 910
Qy 841 TTGATGAACAGTACATGAGTTTGTGAGTGAAGATTAAGTGTGAACTTGTGATTTCTCA 900
Db 911 TTGATGAACAGTACATGAGTTTGTGAGTGAAGATTAAGTGTGAACTTGTGATTTCTCA 970
Qy 901 AAGGTTTTTTCAGCAGAAAAATCCTTTTGAATTTATGAAAAACATTTCTTTAGAAAGAAA 960
Db 971 AAGGTTTTTTCAGCAGAAAAATCCTTTTGAATTTATGAAAAACATTTCTTTAGAAAGAAA 1030
Qy 961 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1020
Db 1031 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1090
Qy 1021 ACAGATAAGCTCTTACCTTTGGATGAGATTTT 1053
Db 1091 ACAGATAAGCTCTTACCTTTGGATGAGATTTT 1123

LOCUS CR627376 3501 bp mRNA linear HTC 22-SEP-2004
DEFINITION Homo sapiens mRNA; cDNA DKFZp686M05248 (from clone DKFZp686M05248).
ACCESSION CR627376
VERSION CR627376.1 GI:50949847
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3501)
AUTHORS Otenwaelder,B., Obermaier,B., Deutschenbaier,S., Schaiipp,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
CONSTRM The German cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix (Martinsried/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp686M05248) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp686M05248
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.
FEATURES
source Location/Qualifiers
1..3501
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686M05248"
/tissue_type="salivary gland"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
/note="ribonucleotide reductase M2 B (TP53 inducible)"
gene 1..3501
/gene="DKFZp686M05248"
332..1225
/gene="DKFZp686M05248"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH10473.1"
/db_xref="GI:50949848"
/translation="MYKQAOAFTAEVDLSKDLPHNKLKADKVFISHILAFFAA SDGVNENLVERPSQVQVPEARCFYQILLENHSEMYSLLDITRIDPKGRFLF NATEWPKKKADWALRWIADKSTFGERVVAFVAVGVFSGFPAIFWLKRGRLM PGLTFSNELISRDEGLHCDPACLMFOYLVNKPPEERVREIIVDAVKIQEFLTEALPV GLIGMNCILMKQYIEFVADRLLVELGFSKVFQAEENPDFMENISLEGKTNFERRVSE YORFAVMAETTDNVFTLDADF"
ORIGIN
Query Match 95.8%; Score 1009; DB 3; Length 3501;
Best Local Similarity 98.5%; Pred. No. 2.9e-257;
Matches 1018; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
Qy 21 GGAAGCGCGCGGCTGGATGAGATGAGAGATCATCTTCAGACACCAAGTGAAT 80
Db 190 GGACGAGCTTGGAGGCTCAGGCCGGGAGATCATCTTCAGACACCAAGTGAAT 249
Qy 81 AAGTCAATGAAGAGCACTCTTAAGAAAGAGTTCTGCCCGGTTTGTCTTCCAAT 140
Db 250 AAGTCAATGAAGAGCACTCTTAAGAAAGAGTTCTGCCCGGTTTGTCTTCCAAT 309
Qy 141 CCAGTACCCTGATATTGGAAATGTATAACAGGCACAGCTTCCTTCGACAGAGA 200
Db 310 CCAGTACCCTGATATTGGAAATGTATAACAGGCACAGCTTCCTTCGACAGAGA 369
Qy 201 AGAGGTGACTTATCAAGGATCTCCCTCACTGGAACAAAGCTTAAAGCAGATGAGA 260
Db 370 AGAGGTGACTTATCAAGGATCTCCCTCACTGGAACAAAGCTTAAAGCAGATGAGA 429


```
QY 261 CTTTCATCTCTCATCTTACATCTTACGCTTTTTCGAGCCAGTGTGGAATGTAAATGAAATTT 320
Db 430 CTTTCATCTCTCATCTTACGCTTTTTCGAGCCAGTGTGGAATGTAAATGAAATTT 489
QY 321 GGTGGAGCCCTTTAGTCAGAGGTGAGGTTCCAGAGGCTCGCTGTTTCTATGCGCTTCA 380
Db 490 GGTGGAGCCCTTTAGTCAGAGGTGAGGTTCCAGAGGCTCGCTGTTTCTATGCGCTTCA 549
QY 381 AATTCTCATCAGAGATGTTTCACTCAGAGATGTACAGTTTGTGATAGACACTTACATCAG 440
Db 550 AATTCTCATCAGAGATGTTTCACTCAGAGATGTACAGTTTGTGATAGACACTTACATCAG 609
QY 441 AGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAACCATGCCCTATGTGTAAGAA 500
Db 610 AGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAACCATGCCCTATGTGTAAGAA 669
QY 501 AAAAGCAGATTTGGGCTTCGATGATAGCAGATAGAAAATCTACTTTTGGGAAAGAGT 560
Db 670 AAAAGCAGATTTGGGCTTCGATGATAGCAGATAGAAAATCTACTTTTGGGAAAGAGT 729
QY 561 GGTGGCCTTTGCTGTGATAGAGGATTTTCTTCAGGATCTTTTGTGCTGCTATATCTG 620
Db 730 GGTGGCCTTTGCTGTGATAGAGGATTTTCTTCAGGATCTTTTGTGCTGCTATATCTG 789
QY 621 GCTTAAAGAGAGAGGTCTTTATGCCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAG 680
Db 790 GCTTAAAGAGAGAGGTCTTTATGCCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAG 849
QY 681 TGAAGGACTTCACTGTGACTTTTGTGCTGCTGATGTTTCCAAATGAACTCATCAGCAGAG 740
Db 850 TGAAGGACTTCACTGTGACTTTTGTGCTGCTGATGTTTCCAAATGAACTCATCAGCAGAG 909
QY 741 AGAAGAAAGGTCAGGAGATCAITGTTGATGCTGTCAAAATGACAGAGATTTTAAAC 800
Db 910 AGAAGAAAGGTCAGGAGATCAITGTTGATGCTGTCAAAATGACAGAGATTTTAAAC 969
QY 801 AGAGCCTTGCAGTGGCTCATTTGGATGAATGCAATTTTGAAGAAACAGTACATGGA 860
Db 970 AGAGCCTTGCAGTGGCTCATTTGGATGAATGCAATTTTGAAGAAACAGTACATGGA 1029
QY 861 GTTGTAGCTGACAGATCTTGTGAACTTGGATCTCAAGTCTTCAAGTCTTTCAGGAGAAA 920
Db 1030 GTTGTAGCTGACAGATCTTGTGAACTTGGATCTCAAGTCTTTCAGGAGAAA 1089
QY 921 TCCTTTTGAATTTATGAAACATTTCTTTAGAAGGAAAAACAAATTTCTTTGAGAAACG 980
Db 1090 TCCTTTTGAATTTATGAAACATTTCTTTAGAAGGAAAAACAAATTTCTTTGAGAAACG 1149
QY 981 AGTTTCAGAGTATCAGCGTTTTCAGATGATGAGCAGAAACACAGATAAGCTTTCACCTT 1040
Db 1150 AGTTTCAGAGTATCAGCGTTTTCAGATGATGAGCAGAAACACAGATAAGCTTTCACCTT 1209
QY 1041 GGATCGAGATTTT 1053
Db 1210 GGATCGAGATTTT 1222

RESULT 5
LOCUS AY398973
DEFINITION Homo sapiens HCM0069 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY398973
VERSION AY398973.1 GI:39754962
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1056)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
```

```
TITLE
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1056)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..1056
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1056
/locus_tag="HCM0069"
gene
ORIGIN
Query Match 89.6%; Score 943.4; DB 9; Length 1056;
Best Local Similarity 89.6%; Pred. No. 6.2e-240;
Matches 944; Conservative 0; Mismatches 109; Indels 0; Gaps 0;
```

```
QY 1 ATGGGCGACCCGAAAGCGCGGAGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGGCGACCCGAAAGCGCGGAGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACCAACGAAAGTGAATTAAGTCAATTAAGAGAGCCACTCTTAAGAAAGAGTTCGCG 120
Db 61 GACACCAACGAAAGTGAATTAAGTCAATTAAGAGAGCCACTCTTAAGAAAGAGTTCGCG 120
QY 121 CGGTTTGTCTATCTTTCCATCCAGTACCTGATATTTGGAAAATGTATAACAGGCACAG 180
Db 121 CGGTTTGTCTATCTTTCCATCCAGTACCTGATATTTGGAAAATGTATAACAGGCACAG 180
QY 181 GCTTCCTTCTGGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAACAG 240
Db 181 GCTTCCTTCTGGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAACAG 240
QY 241 CTTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT 300
Db 241 CTTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT 300
QY 301 GGAAATGTTAAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
Db 301 GGAAATGTTAAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
QY 361 CGCTGTTTCTATGCGCTTTCAAAATTCATCATGAGAAATGTTTCTCAGAGATGTACAGTTG 420
Db 361 CGCTGTTTCTATGCGCTTTCAAAATTCATCATGAGAAATGTTTCTCAGAGATGTACAGTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTAATGCAATGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTAATGCAATGAA 480
QY 481 ACCATGCCCTATGTTTAAGAAAAGCAGATTTGGGCTTCGCGATGATGAGATAGAAAA 540
Db 481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAAGAGAGTTCCTCTCAGGA 600
Db 541 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 600
QY 601 TCTTTTGTGCTATATCTTGGCTTAAAGAAAGAGAGGCTTATGCGAGGACTCACATTTTCC 660
Db 601 TCTTTTGTGCTATATCTTGGCTTAAAGAAAGAGAGGCTTATGCGAGGACTCACATTTTCC 660
```

```

QY 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
DB 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
QY 721 TACTTAGTAATAAGCCTTCAGAGAAAGGGTCAGGAGATCACTTGTGATGCTGTCAAA 780
DB 721 TACTTAGTAATAAGCCTTCAGAGAAAGGGTCAGGAGATCACTTGTGATGCTGTCAAA 780
QY 781 ATTGACGACGAGTATTTTAAACAGAGCCTTGCACAGTTCAGTGGCTCAATGGAATGAATTCATT 840
DB 781 ATTGACGACGAGTATTTTAAACAGAGCCTTGCACAGTTCAGTGGCTCAATGGAATGAATTCATT 840
QY 841 TTGATGAAACAGTACATGAGTTGTAGCTGAGCAGATTAATCTGTGAACTTGGATTTCTCA 900
DB 841 TTGATGAAACAGTACATGAGTTGTAGCTGAGCAGATTAATCTGTGAACTTGGATTTCTCA 900
QY 901 AAGGTTTTTCAGGCAGAAAATCCTTTTGTATTTATGAGAAACATTTCTTTAGAGAGAAA 960
DB 901 AAGGTTTTTCAGGCAGAAAATCCTTTTGTATTTATGAGAAACATTTCTTTAGAGAGAAA 960
QY 961 ACAAAATTTCTTGAGAAAACAGTTTCAGAGATATCAGCGTTTTCAGTTATGGCAGAAAACC 1020
DB 961 ACAAAATTTCTTGAGAAAACAGTTTCAGAGATATCAGCGTTTTCAGTTATGGCAGAAAACC 1020
QY 1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053
DB 1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053

```

RESULT 6

```

AL547501 1083 bp mRNA linear EST 25-MAR-2004
LOCUS AL547501 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1011YF14 5-PRIME, mRNA sequence.
ACCESSION AL547501
VERSION EST.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:31269332.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with NotI and cloned
into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4436.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0D1011DC070P1&c=4436.r.
Location/Qualifiers
1. .1083
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1011YF14"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized."

```

FEATURES source

ORIGIN

Query Match 88.1%; Score 928; DB 1; Length 1083;

RESULT 7

```

Best Local Similarity 97.9%; Pred. No. 7.9e-236;
Matches 975; Conservative 6; Mismatches 11; Indels 4; Gaps 4;
QY 1 ATGGGCGACCCGGAAGCGGAAAGCGCGG -GCTGGATCAGATGAGAGATCATCTTTC 59
DB 27 ATGGGCGACCCGGAAGCGGAAAGCGCGG -GCTGGATCAGATGAGAGATCATCTTTC 86
QY 60 AGACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTCTTCG 119
DB 87 AGACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTCTTCG 146
QY 120 CCGGTTTGTCACTTTTCCAATCCAGTACCTCTGATATTTGGAAAATGTATATAACAGGACA 179
DB 147 CCGGTTTGTCACTTTTCCAATCCAGTACCTCTGATATTTGGAAAATGTATATAACAGGACA 206
QY 180 GGCCTTCTCTTGACAGCAGCAGAGGTGTGATTTATCAAAAGATCTCCCTCACTGGAACA 239
DB 207 GGCCTTCTCTTGACAGCAGCAGAGGTGTGATTTATCAAAAGATCTCCCTCACTGGAACA 266
QY 240 GCTTAAAGCAGATGAGAAGTACTTCTCATCTCTACATCTTAGCCCTTTTTCAGCCAGTGA 299
DB 267 GCTTAAAGCAGATGAGAAGTACTTCTCATCTCTACATCTTAGCCCTTTTTCAGCCAGTGA 326
QY 300 TGAATTTGTAATGAAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGTTCAGAGGC 359
DB 327 TGAATTTGTAATGAAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGTTCAGAGGC 386
QY 360 TGCCTGTTTCTATGGCTTTCAAATCTCATCTGAGAATGTTTCACTCAGAGATGTACAGTTT 419
DB 387 TGCCTGTTTCTATGGCTTTCAAATCTCATCTGAGAATGTTTCACTCAGAGATGTACAGTTT 446
QY 420 GCTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAATTTTATTTAATGCAATTGA 479
DB 447 GCTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAATTTTATTTAATGCAATTGA 506
QY 480 AACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTCGCGATGAGTAGCAGATAGAAA 539
DB 507 AACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTCGCGATGAGTAGCAGATAGAAA 566
QY 540 ATCTACTTTTGGGAAAGAGTGTGGCTTTCGCTGTAGAGAGAGTTTCTTCTCAGG 599
DB 567 ATCTACTTTTGGGAAAGAGTGTGGCTTTCGCTGTAGAGAGAGTTTCTTCTCAGG 626
QY 600 ATCTTTTGTCTGTATATTTCTGGCTTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTC 659
DB 627 ATCTTTTGTCTGTATATTTCTGGCTTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTC 686
QY 660 CAATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTTGCCTGATGTTCCA 719
DB 687 CAATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTTGCCTGATGTTCCA 746
QY 720 ATACTTAGTAATAAGCCTTCAGAGAAAGGGTCAGGAGATCACTTGTGATGCTGTCAA 779
DB 747 ATACTTAGTAATAAGCCTTCAGAGAAAGGGTCAGGAGATCACTTGTGATGCTGTCAA 806
QY 780 AATTGACGAGGAGTATTTTAAACAGAGCCTTGCAGATTGGCTCATTCGAAATGAATTGCAT 839
DB 807 AATTGACGAGGAGTATTTTAAACAGAGCCTTGCAGATTGGCTCATTCGAAATGAATTGCAT 866
QY 840 TTTGATGAAACAGTACATTTAGTGTGTAGCTGACAGATTAATTG -TGGAACTTGGATTTCT 898
DB 867 TTTGATGAAACAGTACATTTAGTGTGTAGCTGACAGATTAATTG -TGGAACTTGGATTTCT 926
QY 899 CAAAGGTTTTTCAGGAGAGAAATCCCTTTTGAATTTTATGAAAAACATTTCTTTAGAGAA 958
DB 927 CAAAGGTTTTTCAGGAGAGAAATCCCTTTTGAATTTTATGG -AAAMAATTTCTTTAGAG -A 984
QY 959 AAACAATTTCTTTGAGAAACAGAGTTTCAGAGATTC 994
DB 985 AAAMAATTTCTTTKAGAACGAGTTTTRAGWWCAGC 1020

```

```

AY398974      1008 bp      DNA      linear      GSS 12-DEC-2003
LOCUS      Pan troglodytes HCM0069 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION      genomic survey sequence.
ACCESSION      AY398974
VERSION      AY398974.1 GI:39754963
KEYWORDS      GSS.
SOURCE      Pan troglodytes (chimpanzee)
ORGANISM      Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE      1 (bases 1 to 1008)
AUTHORS      Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
      Todd,M.A., Tanenbaum,D.M., Civeillo,D.R., Lu,F., Murphy,B.,
      Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
      Adams,M.D. and Cargill,M.
TITLE      Inferring nonneutral evolution from human-chimp-mouse orthologous
      gene trios
JOURNAL      Science 302 (5652), 1960-1963 (2003)
PUBMED      14671302
REFERENCE      2 (bases 1 to 1008)
AUTHORS      Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
      Todd,M.A., Tanenbaum,D.M., Civeillo,D.R., Lu,F., Murphy,B.,
      Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
      Adams,M.D. and Cargill,M.
TITLE      Direct Submission
JOURNAL      Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
      Rockville, MD 20850, USA
COMMENT      This sequence was made by sequencing genomic exons and ordering
      them based on alignment.
FEATURES             Location/Qualifiers
     source           1..1008
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene             <1..>1008
                     /locus_tag="HCM0069"
ORIGIN
Query Match      78.7%; Score 828.2; DB 9; Length 1008;
Best Local Similarity 82.6%; Prid. No. 3e-209;
Matches 830; Conservative 0; Mismatches 175; Indels 0; Gaps 0;
QY      49 AGATCATCTTCAGACACCAACGAAAGTGAATAAGTCAAAATGAAGAGCCATCTCTAAGA 108
DB      1 AGATCATCTTCAGACACCAACGAAAGTGAATAAGTCAAAATGAAGAGCCATCTNTAAGA 60
QY      109 AAGAGTTCTCGCGGTTTGTTCATCTTCCAAATCCAGTACCCCTGATATTTGGAAAATGTAT 168
DB      61 AAGAGTTCTCGCGGTTTGTTCATCTTCCAAATCCAGTACCCCTGATATTTGGAAAATGTAT 120
QY      169 AAACAGGCACAGGCTTCCTCTTGGACAGCAGAGAGTTGACTTATCAAGGATCTCCCT 228
DB      121 AAACAGGCACAGGCTTCCTCTTGGACAGCAGAGAGTGCAGTATCAAGGATCTCCCT 180
QY      229 CACTGGAACAGCTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTACGCTTTTTT 288
DB      181 CACTGGAACAGCTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTACGCTTTTTT 240
QY      289 GCAGCCAGTGTGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGGCAG 348
DB      241 GCAGCCAGTGTGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGGCAG 300
QY      349 GTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTTCTCATCAGAGAATGTTCACTCAGAG 408
DB      301 GTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTTCTCATCAGAGAATGTTCACTCAGAG 360
QY      409 ATGTACAGTTTGTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTT 468
DB      361 ATGTACAGTTTGTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTT 420
QY      469 AATGCAATTGAAACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGCATGGATA 528

```

```

421 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 480
QY      529 GCAGATAGAAATCTACTTTTGGGGAAGAGTGTGGCTTGTGCTGTAGAGAGAGTT 588
DB      481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540
QY      589 TTCTTCTCAGGATCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGAGTCTTATGCGCAGGA 648
DB      541 TTCTTCTCAGGATCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGAGTCTTATGCGNNNN 600
QY      649 CTCACATTTTCCAAATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGC 708
DB      601 NTCACATTTTCCAAATGAATCAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 660
QY      709 CTCATGTTTCCAAATCTAGTAAATAAGCCCTCAGAAGAAAGGTCAGGAGATCATTTGTT 768
DB      661 CTGATGTTTCCAAATCTAGTAAATAAGCCCTCAGAAGAAAGGTCAGGAGATCATTTGTT 720
QY      769 GATGCTGTCAAAATTTGAGCAGGAGTCTTTTAACAGAAAGCCTTGCAGTTGGCCTCATTTGA 828
DB      721 GATGCTGTCAAAATTTGAGCAGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 780
QY      829 ATCAATTTGCAATTTTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGAA 888
DB      781 ATCAATTTGCAATTTTGATGANACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGAA 840
QY      889 CTTGGATTTCTCAAGGTTTTTTTCAGGCGAGAAATCCTTTTGATTTTATGGAATAAATTTCT 948
DB      841 CTTGGATTTCTCAAGGTTTTTTTCAGGCGAGAAATNNNTTGAATTTTATGGAATAAATTTCT 900
QY      949 TTAGAAGGAAAAACAATTTCTTTTGAGAAAAGGTTTCAGAGATATCAGCGTTTTCAGTT 1008
DB      901 TTAGAAGGAAAAACAATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 960
QY      1009 ATGCGAGAAACCCAGATACGCTTTCACCTTTGATGATGATGATGATGATGATGAT 1053
DB      961 ATGCGAGAAACCCAGATACGCTTTCACCTTTGATGATGATGATGATGATGATGATGAT 1005

```

RESULT 8
BQ441857
LOCUS

DEFINITION BQ441857 879 bp mRNA linear EST 24-MAY-2002
AGENCOURT 7827448 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6014179
5', mRNA sequence.

ACCESSION BQ441857
VERSION BQ441857.1 GI:21180933
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 879)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs@mail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LLAM13208 row: d column: 20
High quality sequence stop: 662.
Location/Qualifiers

FEATURES
source

1..879
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6014179"
/tissue_type="embryonal carcinoma, cell line"

```
/lab host="DH10B (phage-resistant)"
/clone lib="NIH_MGC_92"
/notes="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dr primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 75.2%; Score 791.6; DB 5; Length 879;
Best Local Similarity 98.6%; Pred. No. 1.6e-199;
Matches 820; Conservative 0; Mismatches 9; Indels 3; Gaps 2;

QY 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 60
Db 22 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 81

QY 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Db 82 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 141

QY 121 CGGTTTGTTCATCTTCCAAATCAGTACCTGTATATTGGAAAATGTATATAACAGGCACAG 180
Db 142 CGGTTTGTTCATCTTCCAAATCAGTACCTGTATATTGGAAAATGTATATAACAGGCACAG 201

QY 181 GCTTCTCTCTGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGACACAG 240
Db 202 GCTTCTCTCTGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGACACAG 261

QY 241 CTTAAAGCAGATGAGAAGTACTTCTCTCACATCTTAGCCCTTTTTCAGCAGCTGAT 300
Db 262 CTTAAAGCAGATGAGAAGTACTTCTCTCACATCTTAGCCCTTTTTCAGCAGCTGAT 321

QY 301 GGAATTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db 322 GGAATTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 381

QY 361 CGCTGTTCTATGGCTTTCAAAATCTCATCGAGATGTTTCACTCAGAGATGTACAGTTTG 420
Db 382 CGCTGTTCTATGGCTTTCAAAATCTCATCGAGATGTTTCACTCAGAGATGTACAGTTTG 441

QY 421 CTGATAGACACTTATACATCAGAGATCCAAAGAAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 442 CTGATAGACACTTATACATCAGAGATCCAAAGAAAAGGGAATTTTATTTAATGCAATTGAA 501

QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATGCGGCTTCGGATGGATAGCAGATAGAAA 540
Db 502 ACCATGCCCTATGTTAAGAAAAGCAGATGCGGCTTCGGATGGATAGCAGATAGAAA 561

QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGGAGTTTCTCTCAGGA 600
Db 562 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGGAGTTTCTCTCAGGA 621

QY 601 TCTTTTGTCTATATCTGGCTAAAGAGAGAGGTCTTATGCGAGACTCACTTTTTC 660
Db 622 TCTTTTGTCTATATCTGGCTAAAGAGAGAGGTCTTATGCGAGACTCACTTTTTC 681

QY 661 AATGAATCATCAGCAGAGATCAAGGACTTCACTGTGACTTTTCTGCTGCTGATGTTCCAA 720
Db 682 AATGAATCATCAGCAGAGATCAAGGACTTCACTGTGACTTTTCTGCTGCTGATGTTCCAA 741

QY 721 TACTTAGTAAATAGGCTTTCAGAGAAAGGG--TCAGGGAGATCATTTGATGCTGTCA 778
Db 742 TACTTAGTAAATAGGCTTTCAGAGAAAGGGGTCAGGGGAGATCATTTGATGCTGTCA 801

QY 779 AATTT-GAGCAGAGTTTTTAAACAGAGCCTTGCCAGTTGGCTCAATCGAA 829
Db 802 AATTTGAGCAGAGTTTTTAAACAGAGCCTTGCCAGTTGGCTCAATTTGGA 853

RESULT 9
AY398975 1043 bp DNA linear GSS 12-DEC-2003
LOCUS
```

```
DEFINITION Mus musculus HCM0069 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY398975
VERSION AY398975.1 GI:39754964
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1043)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrieres,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1043)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrieres,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..1043
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1043
/locus_tag="HCM0069"
ORIGIN
Query Match 73.9%; Score 778.6; DB 9; Length 1043;
Best Local Similarity 81.3%; Pred. No. 5e-196;
Matches 855; Conservative 0; Mismatches 184; Indels 13; Gaps 1;

QY 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
Db 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGAGCGCGGAGGAGGTGACGCAATTGTTGCG 60

QY 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGTGTGAGGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCCCGA 120

QY 121 CGGTTTGTTCATCTTCCAAATCCAGTACCTGTATTTGGAAAATGTATATAACAGGCACAG 180
Db 121 CGGTTTGTTCATCTTCCAAATCCAGTATCTGTATATCTGGAGAAATGTACAAGCAAGCAG 180

QY 181 GCTTCTCTCTGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
Db 181 GCATCTTTCTGGACAGCAGAGAGGTGACTTGTCAAGAGCCTCCCTCACTGGAACAAG 240

QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTCACTTTAGCCCTTTTTCAGCAGCTGAT 300
Db 241 CTTAAATCCGACGAGAAGTATTTTATCTCCCACTCTTAGCCCTTTTTCAGCAGCTGAT 300

QY 301 GGAATTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db 301 GGAATTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360

QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGATGTTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAAGATTTCTCATCGAGATGTTTCACTCAGAGATGTACAGTTTA 420

QY 421 CTGATAGACACTTATACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 421 CTAATAGACACTTATACATCAGAGATCCCAAGAAAAGNNNNNNNNNNNNNNNNNNNNNN 480
```

```
QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATGTGGCCCTTGCATGGATAGCAGATAGAAAA 540
Db 481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTTCCTCTCAGGA 600
Db 541 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 600
QY 601 TCTTTTGTCTATATCTTGGCTTAAAGAGAGAGGCTTATGCCAGGACCTACCTTTTTC 660
Db 601 TCAATTTGTGCAATATCTTGGCTTAAAGAGAGAGGCTCATGCTGGACATGATTTTCA 660
QY 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCCCTGCTGATGTTCAA 720
Db 661 AATGAACATCATCAGCAGGAGATGAAGACTTCACTGTGACTTTGCCCTGCTGATGTTCAA 720
QY 721 TACTTAGTAATAAGCCCTTCAGAGAAAGGCTCAGGAGATCAATTTGTGATGCTGTCAA 780
Db 721 TACTTGGTAAACAAGCCCTTCAGAGAGATAGAGTGAAGGAAATCAATGCTGATGCTGTCAA 780
QY 781 ATTGACGAGGAGTTTTTAACAGAGAGCTTGCAGTGTGGCCCTCATTTGGAATGAAATTCATT 840
Db 781 ATCGAG-----CAGAGCCCTTGGCTTGGCCCTCATTTGGAATGAAATTCATT 840
QY 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATCTGTGAACTTTGATTTCTCA 900
Db 828 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATCTGTGAACTTTGATTTCTCA 887
QY 901 AAGTTTTTTCAGCAGAAAAATCCTTTTGAATTTTATGAAACAAATTTCTTTAGAGGAAAA 960
Db 888 AAGATTTTTCAGCAGAAAAATCCTTTGATTTTATGAGAAACAAATTTCTTTAGAGGAAAA 947
QY 961 ACAATTTTTCAGAGAAAGAGTTTCAGAGATATCAGGTTTTCAGTATGCGAGAAACC 1020
Db 948 ACAATTTTTCAGAGAAAGAGTTTCTGAGTATCAGCGATTTTCAGTATGCGCGAAACC 1007
QY 1021 ACAGATAACGCTTTCACCTTTGGATGAGATTT 1052
Db 1008 ACAGATATGCTTTCACCTTTGATGAGATTT 1039
```

```
RESULT 10
LOCUS CN163214/c 892 bp mRNA linear EST 02-APR-2004
DEFINITION 952895 MARC 4P1G Sus scrofa cDNA 3', mRNA sequence.
ACCESSION CN163214
VERSION CN163214.1 GI:46177644
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 892)
Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A.,
Noneman,D.J., Wray,J.E. and Keele,J.W.
Porcine EST collection using a normalized library constructed from
embryos representing early developmental stages
Unpublished (2003)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: TM8060 row: G column: 4
Seq primer: TAGAAGGCACATCGAGG.
1..892
Location/Qualifiers
/organism="Sus scrofa"
/mol_type="mRNA"
```

FEATURES

source

```
/db xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 4P1G"
/notes="vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with combined RNA from day-10, day-13,
day-15, day-25, and day-30 whole embryos."
```

ORIGIN

```
Query Match 69.4%; Score 731.2; DB 7; Length 892;
Best Local Similarity 94.1%; Pred. No. 2e-183;
Matches 760; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 246 AGCAGATGAGAAGTACTTCTCATCTCTAGCCCTTTTTCAGCCAGTGTGGAAT 305
Db 892 ATCAGATGAGAAGTATTTATCTCTCATCTTAGCCCTTTTTCAGCCAGTGTGGAAT 833
QY 306 TGTAAATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGCGCTGCTG 365
Db 832 TGTGAATGAAAACTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGCGCGCTG 773
QY 366 TTCTATGGCTTTCAAAATCTCATCGAATGTTTCACTCAGAGATGTACAGTTTGTGAT 425
Db 772 TTCTACGGCTTTCAAAATCTCATCGAATGTTTCACTCAGAGATGTACAGTTTGTGAT 713
QY 426 AGACACTTATCATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAAACCAT 485
Db 712 AGACACTTATCATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAAACCAT 653
QY 486 GCCCTATGTTAAAGAAAAAGCAGATTTGGCCCTTGCATGGATAGCAGATAGAAAACTTAC 545
Db 652 GCCATATGTTAAAGAAAAAGCAGATTTGGCCCTTGCATGGATAGCAGATAGAAAACTTAC 593
QY 546 TTTTGGGAAAGAGTGTGGCTTTGCTGTGTAGAGAGGTTTCTCTCAGGATCTTT 605
Db 592 TTTTGGGAAAGAGTGTGGCTTTGCTGTGTAGAGAGGTTTCTCTCGGATCTTT 533
QY 606 TGCTGCTATATTCTGGCTAAAGAGAGAGGTTTATGCCAGGACTCCTCTTTTCCAAATGA 665
Db 532 TGCTGCTATATTCTGGTTAAAGAAAAAGAGGTTTATGCCCTGGACTCAGCTTTTCCAAATGA 473
QY 666 ACTCATCAGCAGAGATCAAGGACTTCACTGTGACTTTTGTGCTTGCCTGTATGTTTCAATCTT 725
Db 472 ACTCATCAGCAGAGATCAAGGAGGCTTCACTGTGACTTTTGTGCTTGCCTGTATGTTTCAATCTT 413
QY 726 AGTAAATTAAGCCTTTCAGAGAAAGGGTCAGGAGATCATTTGTGTGCTGTCAAAATGA 785
Db 412 GSTAAATTAAGCCTTTCAGAGCAAGAGTTAGGAGATCATTTGTGTGCTGTGAAATGA 353
QY 786 GCAGGAGTTTAAACAGAGCGTTGCCAGTTGGCTCATTTGGAATGAAATGCAATTTGAT 845
Db 352 GCAGGAGTTTAAACAGAGCGTTGCCAGTTGGCTCATTTGGAATGAAATTTGTGTTTGTAT 293
QY 846 GAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGATTTCTCAAGCT 905
Db 292 GAAACAGTATATTCAGTTTGTAGCTGACAGATTCCTTACAGAACTTTGATTTCTCAAGCT 233
QY 906 TTTTTCAGGCAAAAAATCCTTTTGAATTTTATGAAAAACAATTTCTTTAGAGGAAAAACAAA 965
Db 232 TTTTTCAGGCAAAAAATCCTTTTGAATTTTATGAAAAACAATTTCTTTAGAGGAAAAACAAA 173
QY 966 TTTTCTTTGAGAAAAGAGTTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACACAGA 1025
Db 172 TTTTCTTTGAGAAAAGAGTTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACACAGA 113
QY 1026 TAACGCTTTCACCTTCGATGCAGATTTT 1053
Db 112 TAATGCTTTCACCTTCGATGCAGATTTT 85
```

RESULT 11

CF995079
LOCUS

CF995079 914 bp mRNA linear EST 25-NOV-2003

DEFINITION AGENCOURT 15621478 NIH MGC 147 Homo sapiens cDNA clone
IMAGE:30528806 5', mRNA sequence.
ACCESSION CF995079
VERSION CF995079.1 GI:38511139
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 914)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NDAM615 row: f column: 15
High quality sequence stop: 677.
Location/Qualifiers
1..914
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30528806"
/tissue_type="Human Placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH MGC 147"
/note="Organ: Placenta; Vector: pBluescriptR; Site_1:
all-XhoI; Site_2: BamH; Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: This is
a NIH_MGC library."

ORIGIN
Query Match 69.0%; Score 726.6; DB 7; Length 914;
Best Local Similarity 99.0%; Pred. No. 3.4e-182;
Matches 763; Conservative 0; Mismatches 4; Indels 4; Gaps 3;
QY 1 ATGGGCGACCCGGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 87 ATGGGCGACCCGGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 146
QY 61 GACACCAACGAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAGAGTTCTCGC 120
DB 147 GACACCAACGAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAGAGTTCTCGC 206
QY 121 CGTTTGTTCATCTTCCATCCAGTACCTGTATTTGGAAAATGTATAACAGGCACAG 180
DB 207 CGTTTGTTCATCTTCCATCCAGTACCTGTATTTGGAAAATGTATAACAGGCACAG 266
QY 181 GCTTCTCTTGACAGCAGAGAGGTGTGACTTATCAAAGGATCTCCCTCACTGGAACAAG 240
DB 267 GCTTCTCTTGACAGCAGAGAGGTGTGACTTATCAAAGGATCTCCCTCACTGGAACAAG 326
QY 241 CTTAAGCAGATGAGAAGTACTTCATCTCTCAATCTTAGCCCTTTTTCAGCCAGTAT 300
DB 327 CTTAAGCAGATGAGAAGTACTTCATCTCTCAATCTTAGCCCTTTTTCAGCCAGTAT 386
QY 301 GGAATTGTAAATGAAAATTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360

387 GGAATTGTAAATGAAAATTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 446
361 CGCTGTTTCTATGGCTTCAAAATTCATCAGAGATGTTCTACTCAGAGATCTCAGATTG 420
447 CGCTGTTTCTATGGCTTCAAAATTCATCAGAGATGTTCTACTCAGAGATCTCAGATTG 506
421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTAATGCAATTGAA 480
507 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTAATGCAATTGAA 566
481 ACCATGCCCTATGTTAAGAAAAGCAGATGGGCCCTTGGATGGATGAGATAGAGAAA 540
567 ACCATGCCCTATGTTAAG-AAAAGCAGATGGGCCCTTGGATGGATGAGATAGAGAAA 624
541 TCTACTTTTGGGGAAGAGTGGTGGCTTGTCTGTAGAGGAGTTTCTTCTCAGGA 600
625 TCTACTTTTGGGGAAGAGTGGTGGCTTGTCTGTAGAGGAGTTTCTTCTCAGGA 684
601 TCTTTTGTCTATATCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACATTTTTC 660
685 TCTTTTGTCTATATCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACATTTTTC 744
661 AATGAACATCATCAGCAGAGATGAAGACTTCTACTGTGCTTGTCTGCTGATGTTCAA 720
745 AATGAACATCATCAGCAGAGATGAAGACTTCTACTGTGCTTGTCTGCTGATGTTCAA 804
721 TACTTAGTAATAAGCC-TTCAGAGAAAGGCTC-AGGAGATCATTTGTTG 769
805 TACTTAGTAATAAGCCTTTCAAAAGAAAGGGTCTAGGAAGATCAITGTTG 855

RESULT 12
LOCUS CO648380
DEFINITION ILLUMIGEN MQC 41729 Katze MMPB2 Macaca mulatta cDNA clone
ILLUM:25149 5' similar to Bases 5 to 770 highly similar to human
RRM2B (Hs.512592), mRNA sequence.
ACCESSION CO648380
VERSION CO648380.1 GI:50569874
KEYWORDS EST.
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecoidea; Macaca.
REFERENCE 1 (bases 1 to 870)
AUTHORS Katze, M.G., Thomas, M., Korth, M., Iadonato, S.P. and Magnus, C.L.
TITLE Large-scale Rhesus Macaque cDNA Sequencing
JOURNAL Unpublished (2003)
COMMENT Contact: C. Magnus
Illumigen Biosciences Inc.
2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
Tel: 2063780400
Fax: 2063780408
Email: cmagnus@illumigen.com
Sequenced on 2004.06.02. 691 Q20 bases. Library Preparation: Prof.
Michael Katze Lab at University of Washington DNA Sequencing:
Illumigen Biosciences Inc. For further information, see
<http://www.macaque.org>
PCR Primers
FORWARD: CCCTCACTAAAGGGAACAAA
BACKWARD: CACTATAGGCGGAATTGGTA
Insert length: 870 Std Error: 0.00
Plate: CL000348 row: B column: 06
Seq primer: CCCTCACTAAAGGGAACAAA
POLYA=No. Location/Qualifiers
1..870
/organism="Macaca mulatta"
/mol_type="mRNA"
/strain="Indian"
/db_xref="taxon:9544"


```
/clone="IBIUM:25149"
/sex="male"
/tissue_type="blood"
/cell_type="PBMC"
/dev_stage="adult"
/lab_host="Electromax DH10B"
/clone_lib="Katz MPB2"
/notes="Vector: pDONR 222; Site 1: Berg I; Site 2: BsrG I;
Created from CloneMiner cDNA Library Construction kit
(catalog #18249-029)"

ORIGIN
Query Match 67.5%; Score 710.6; DB 7; Length 870;
Best Local Similarity 98.8%; Pred. No. 6.2e-178;
Matches 716; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 329 GCTTTAGTCAGAGGTCGAGGTCGAGGTCGAGGTCGCTGTTCTATGCGCTTCAAATTCCTCA 388
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 389 TCGAGAAATCTCACTCAGAGATGTCAGTTTCTGTATAGACACTTACATCAGATGATCCCA 448
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 63 TCGAGAAATGTTCAATCAGAGATGTCAGTTTCTGTATAGACACTTACATCAGATGATCCCA 122
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 449 AGAAAGGCGAAATTTTATTAATGCAATGAAACCAATGCGCTATGTTAAGAAAAAGCAG 508
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 123 AGAAAGGCGAAATTTTATTAATGCAATGAAACCAATGCGCTATGTTAAGAAAAAGCAG 182
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 509 ATTGGGCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 568
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 183 ATTGGGCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 242
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 569 TTGCTGCTAGAGAGGATTTTCTCTCAGGATCTTTTCTGCTGATATTTCTGCTGATATTTCTG 628
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 243 TTGCTGCTAGAGAGGATTTTCTCTCAGGATCTTTTCTGCTGATATTTCTGCTGATATTTCTG 302
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 629 AGAGAGGCTTATGCGAGGACTCCTTTTCCAAATGCAATGCAATGCAATGCAATGCAATGCAAT 688
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 303 AGAGAGGCTTATGCGAGGACTCCTTTTCCAAATGCAATGCAATGCAATGCAATGCAATGCAAT 362
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 689 TTCCTGTGACTTTGCTTGCCTGATGTTCCAAATCTTAGTAATAGCCCTTCAGAGAGAA 748
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 749 GGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTTGAGCAGAGGATTTTAAACAGAGGCT 808
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 423 GGGTCAGGAGATCATTTGTTGATGCTGTCCAAATTTGAGCAGAGGATTTTAAACAGAGGCT 482
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 809 TGCCAGTTGGCTCATTTGGAATGAAATGCAATTTGATGAAACAGTACATTTGATGTTGTAG 868
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 483 TGCCAGTTGGCTCATTTGGAATGAAATGCAATTTGATGAAACAGTACATTTGATGTTGTAG 542
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 869 CTGACAGATTACTTTGGAATCTGGAATCTCAAGGTTTTCAGGAGAGAAATCCCTTTTG 928
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 543 CTGACAGATTACTTTGGAATCTGGAATCTCAAGGTTTTCAGGAGAGAAATCCCTTTTG 602
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 929 ATTTTATGAAACATTTCTTTAGAGGAGAAACAAATTTCTTTGAGAACCGATTTTCAG 988
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 603 ATTTTATGAAACATTTCTTTAGAGGAGAAACAAATTTCTTTGAGAACCGATTTTCAG 662
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 989 AGTATCAGGTTTTCAGTTATGGCAGAAACACAGATACGTTCTTACCTTGGATGCGAG 1048
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 663 AGTATCAGGTTTTCAGTTATGGCAGAAACACAGATACGTTCTTACCTTGGATGCGAG 722
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1049 ATTTT 1053
Db | | | | |
QY 723 ATTTT 727
Db | | | | |
```

```
RESULT 13
CD656661
LOCUS
DEFINITION AGENCOURT_14555875 NIA Human H1 Embryonic Stem Cell cDNA Library
CD656661 804 bp mRNA linear EST 18-JUN-2003
Query Match 64.8%; Score 682.8; DB 6; Length 804;
Best Local Similarity 97.3%; Pred. No. 1.5e-170;
Matches 716; Conservative 0; Mismatches 17; Indels 3; Gaps 2;
```

```
ACCESSION CD656661
VERSION CD656661.1
KEYWORDS GI:31897180
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 804)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabs-x@mail.nih.gov
Tissue Procurement: Irene Ginis and Mahendra Rao, NIA
cDNA Library Preparation: Yulan Piao and Minoru Ko
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC c lone distribution information
can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAMS14 row: h column: 17
High quality sequence stop: 680.
FEATURES
location/Qualifiers
1..804
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30427288"
/tissue_type="Embryonic Stem cells"
/cell_line="WA01"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)"
/notes="Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI;
This is a long-transcript enriched cDNA library (Genome
Res. 11: 1553-1558 (2001). [PMID: 1154199]) from WA01
cell line. Undifferentiated human ES cell line WA01/H1
was obtained from WiCell Research Institute, Inc.,
Madison, WI, cultured according to their instructions, on
MEF feeders. They formed round colonies with defined edges
and were positive for alkaline phosphatase, SSEA-4, OCT3,
OCT4, REX1, UTF, TERT, SOX2, CX43 and CX45. They are
negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1,
TUBB3, NES, GFAP, and BOMES. When confluent (18-10 days
after plating), the ES cells from 4 X 6cm dishes were
treated with 1 mg/ml collagenase, type IV
(Invitrogen/GIBCO) for 5-10 min and gently scraped off
with 5 ml pipette. RNA was purified with Trizol Reagent
from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558
(2001). [PMID:1154199]) Double-stranded cDNAs were
synthesized with an Oligo(dT) primer [Invitrogen:
5'-pGACTAGTTCGAGCGCGCCCTTTTCTTTTCTTTT-3'] from
3.4g of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to Lona-linker LL-Sal4, purified by phenol/chloroform
extraction, and separated from free linkers by
Centricon-100 column. Then, the cDNAs were amplified by
long-range high fidelity PCR using Ex Taq polymerase
(Takara) with a primer Sal4-S for 25 cycles. The products
were purified by phenol/chloroform extraction and
Centricon-100 column. The cDNAs were digested with SalI
and NotI enzymes and cloned into SalI/NotI site of
pCMV-SpORf6 plasmid vector. The average insert size is
about 3.6kb."
```

212 TATCAAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTCTATCTCTC 271
1 TATCAAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTCTATCTCTC 60
272 ACATCTTAGC-CTTTTTCAGCCAGTGTGAATTTGAATGAAATTTGGTGGAGCGC 330
61 ACATCTTAGCCTTTTTCAGCCAGTGTGAATTTGAATGAAATTTGGTGGAGCGC 120
331 TTTAGTCAGAGGTGAGGTTCAGAGGCTCGCTGTTCTATGGCTTCAAAATTTCTCATC 390
121 TTTAGTCAGAGGTGAGGTTCAGAGGCTCGCTGTTCTATGGCTTCAAAATTTCTCATC 180
391 GAGATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCAAG 450
181 GAGATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCAAG 240
451 AAAAGGGAATTTTATTAATGCAATGAAACCATGCCCCTATGTTAAGAAAAAGCAGAT 510
241 AAAAGGGAATTTTATTAATGCAATGAAACCATGCCCCTATGTTAAGAAAAAGCAGAT 300
511 TGGGCTTCGGATGGATAGCAGATGAAATCTACTTTTGGGGAAGAGTGGTGGCTTT 570
301 TGGGCTTCGGATGGATAGCAGATGAAATCTACTTTTGGGGAAGAGTGGTGGCTTT 360
571 GCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTGTCTGCTATATTTCTGGCTAAAGAG 630
361 GCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTGTCTGCTATATTTCTGGCTAAAGAG 420
631 AGAGGTCTTATGCAAGGATCTACTTTTCCAATGAACTCATCAGCAGAGATGAAGACTT 690
421 AGAGGTCTTATGCAAGGATCTACTTTTCCAATGAACTCATCAGCAGAGATGAAGACTT 480
691 CACTGTGACTTGTGCTGCTGATGTTTCCAATGAACTCATCAGCAGAGATGAAGACTT 750
481 CACTGTGACTTGTGCTGCTGATGTTTCCAATGAACTCATCAGCAGAGATGAAGACTT 540
751 GTCAAGGAGATCAATTTGCTGCTGCTGCTCAAAATTTGAGCAGAGTTTTAAACAGAGCTTG 810
541 GTCAAGGAGATCAATTTGCTGCTGCTGCTCAAAATTTGAGCAGAGTTTTAAACAGAGCTTG 600
811 CCAGTTGGCTCATTTGAATGAATTTGATTTGATGAAACAGTACATTTGATTTTGTAGCT 870
601 CCAGTTGGCTCATTTGAATGAATTTGATTTGATGAAACAGTACATTTGATTTTGTAGCT 660
871 GACAGATTTACTTGTGAATTTGATTTCTCAAGGTTTTCAGCAGAGAAATTCCTTTTGTAT 930
661 GACAGATTTACTTGTGAATTTGATTTCTCAAGGTTTTCAGCAGAGAAATTCCTTTTGTAT 718
931 TTTATGGAACAACTTT 946
719 TATGGAACAACTTTT 734

RESULT 14
BM801298
LOCUS
DEFINITION BM801298 1061 bp mRNA linear EST 05-MAR-2002
5', mRNA sequence. AGENCOURT_6422860 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5586693
ACCESSION BM801298
VERSION BM801298.1 GI:19118121
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1061)
NIH-MGC <http://mgs.nhl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1AM12354 row: p column: 22
High quality sequence stop: 631.

FEATURES
source

1..1061
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5586693"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/notes="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: Sali; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 62.1%; Score 653.8; DB 4; Length 1061;
Best Local Similarity 97.4%; Pred. No. 8.9e+163;
Matches 717; Conservative 0; Mismatches 13; Indels 6; Gaps 5;
QY 1 ATGGGCGACCCGGAAGGCGCGGCGGCTGATCAGGATGAGAGATCATCTTCA 60
DB 85 ATGGGCGACCCGGAAGGCGCGGCGGCTGATCAGGATGAGAGATCATCTTCA 144
QY 61 GACCAACGAAAGTGAATTAAGTCAAAATGAGAGCCACTCTTAAAGAAAGTCTCGC 120
DB 145 GACCAACGAAAGTGAATTAAGTCAAAATGAGAGCCACTCTTAAAGAAAGTCTCGC 204
QY 121 CGGTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAAATGTATAAACAGGCACAG 180
DB 205 CGGTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAAATGTATAAACAGGCACAG 264
QY 181 GCTTCTCTTGGACAGCAGAGAGTTGATTTATCAAAAGATCTCCCTCACTGGAACAAAG 240
DB 265 GCTTCTCTTGGACAGCAGAGAGTTGATTTATCAAAAGATCTCCCTCACTGGAACAAAG 324
QY 241 CTTAAAGCAGATCAGAGTACTTCACTCTCACATCTTAGCCCTTTTGGAGCCAGTAT 300
DB 325 CTTAAAGCAGATCAGAGTACTTCACTCTCACATCTTAGCCCTTTTGGAGCCAGTAT 384
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGCAGGTTCCAGAGCT 360
DB 385 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGCAGGTTCCAGAGCT 444
QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTATCAGAAATGTTTCACTCAGAGATGTACAGTTTG 420
DB 445 CGCTGTTTCTATGGCTTTCAAATTTCTATCAGAAATGTTTCACTCAGAGATGTACAGTTTG 504
QY 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGAA 480
DB 505 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGAA 564
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTGGATGGATGAGATAGAAAA 540
DB 565 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTGGATGGATGAGATAGAAAA 624
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTAGAGAGGAGTTTCTTCTCAGGA 600
DB 625 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTAGAGAGGAGTTTCTTCTCAGGA 684
QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGTCTTATGCGAGGACTCACTTTTTC 660
DB 685 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTC-TATGCCAGGACTCACTTTTTC 743
QY 661 AATGAACATCAT-CAGCAGAGAT-GAAGGACTTCACTGCTGA--CTTTGCTTGCCTGAT-GT 715

```

Db      744 CATGGACTCATCCACGAGATGGAGGACTTCCTGTGAACCTTTCCTTGCCTGTAGTGT 803
Qy      716 TCCAATACCTTAGTAAA 731
Db      804 TCCAATACCTNAAGAAA 819

RESULT 15
BUI70979
LOCUS   BUI70979
DEFINITION BUI70979 896 bp mRNA linear EST 04-SEP-2002
5', mRNA sequence.
AGENCOURT_7940338 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6143754
VERSION BUI70979
KEYWORDS BUI70979.1 GI:22684963
SOURCE   EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 896)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLMW13467 row: k column: 19
High quality sequence stop: 619.

FEATURES
source
1..896
Location/Qualifiers
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:6143754"
    /tissue_type="retinoblastoma"
    /lab_host="DH10B (phage-resistant)"
    /clone_lib="NIH_MGC_67"
    /note="Organ: eye; Vector: pCMV-SPORT6; Site: 1: NotI;
    Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
    Average insert size 1.75 kb. Library constructed by Life
    Technologies."

ORIGIN
Query Match 62.0%; Score 653.2; DB 5; Length 896;
Best Local Similarity 98.5%; Pred. No. 1.2e-162;
Matches 670; Conservative 0; Mismatches 8; Indels 2; Gaps 1;

Qy      1 ATGGGGACCCCGAAGCGCGGAGCGGCGGCTGGATCAGATCAGAGATCATCTTCA 60
Db      55 ATGGGGACCCCGAAGCGCGGAGCGGCGGCTGGATCAGATCAGAGATCATCTTCA 114
Qy      61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTCTCTCG 120
Db      115 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTCTCTCG 174
Qy      121 CGGTTTGTATCTTCCATCCAGTACCCTGTATTTGGAAAATGTATAAACAGGCACAG 180
Db      175 CGGTTTGTATCTTCCATCCAGTACCCTGTATTTGGAAAATGTATAAACAGGCACAG 234
Qy      181 GCTTCTCTTGACAGCAGAGAGGTGTGACTTATCAAGGATCTCCCTCACTGGAAACAG 240
Db      235 GCTTCTCTTGACAGCAGAGAGGTGTGACTTATCAAGGATCTCCCTCACTGGAAACAG 294
Qy      241 CTTAAAGCAGATGAGAAGTACTTCTCTCTCATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db      295 CTTAAAGCAGATGAGAAGTACTTCTCTCTCATCTTAGCCCTTTTTCAGCCAGTGAT 354

```

```

Qy      301 GGAATTGTAATAAGAAAATTTGGTGGAGCGCTTTAGTCAGAGAGGTGCAGGTTCCAGAGGCT 360
Db      355 GGAATTGTAATAAGAAAATTTGGTGGAGCGCTTTAGTCAGAGAGGTGCAGGTTCCAGAGGCT 414
Qy      361 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCAGAAATGTTCATCAGAGATGTACAGTTTG 420
Db      415 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCAGAAATGTTCATCAGAGATGTACAGTTTG 474
Qy      421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 480
Db      475 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 534
Qy      481 ACCATGCCCTTATGTTAAGAAAAGAGATTTGGGCCCTTTCGATGGATGACAGATAGAAAA 540
Db      535 ACCATGCCCTTATGTTAAGAAAAGAGATTTGGGCCCTTTCGATGGATGACAGATAGAAAA 594
Qy      541 TCTACTTTTGGGAAAAGAGTGGTGGCCTTTTCTCTCTGTAGAAAGAGTTTCTTCTCAGGA 600
Db      595 TCTACTTTTGGGAAAAGAGTGGTGGCCTTTTCTCTCTGTAGAAAGAGTTTCTTCTCAGGA 654
Qy      601 TCTTTTGTCTGTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 658
Db      655 TCTTTTGTCTGTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 714
Qy      659 CCAATGAACCTCATCAGCAGA 678
Db      715 CAATGAACCTCATCAGCAGA 734

```

Search completed: October 30, 2005, 03:23:17
Job time : 2721.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2005, 23:30:35 ; Search time 146 Seconds

(without alignments)
11801.366 Million cell updates/sec

Title: US-10-698-228-12

Perfect score: 1053

Sequence: 1 atggcgaccgacccgaaaggcc.....tcaccttgatgcagatttt 1053

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata1/ina/5A COMB.seq.*
- 2: /cgn2_6/prodata1/ina/5B COMB.seq.*
- 3: /cgn2_6/prodata1/ina/6A COMB.seq.*
- 4: /cgn2_6/prodata1/ina/6B COMB.seq.*
- 5: /cgn2_6/prodata1/ina/6C COMB.seq.*
- 6: /cgn2_6/prodata1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1053	100.0	1053	4	US-10-019-733-12
2	1053	100.0	1081	4	US-10-019-733-4
3	1051.4	99.8	1053	4	US-10-019-733-2
4	1051.4	99.8	4955	4	US-10-019-733-3
5	580.6	55.1	2500	4	US-09-962-665-9
6	580.6	55.1	2500	4	US-09-963-333-9
7	580.6	55.1	2500	4	US-09-962-677-9
8	580	55.1	2500	4	US-09-023-655-1370
9	580	55.1	2500	4	US-09-949-016-145
10	578.4	54.9	2479	4	US-09-949-016-2025
11	439.4	41.7	481	3	US-08-905-223-125
12	429.8	40.8	14176	1	US-08-307-499-1
13	429.8	40.8	14176	1	US-08-307-499-14
14	429.8	40.8	14176	3	US-09-299-268-1
15	429.8	40.8	14176	3	US-09-299-268-14
16	384.8	36.5	1236	4	US-09-248-796A-3914
17	382.8	36.4	825	1	US-08-307-499-40
18	382.8	36.4	825	3	US-09-299-268-40
19	339.4	32.2	1112	1	US-08-136-743B-1
20	317.6	30.2	1083	4	US-09-248-796A-3915
21	316.6	30.1	419	3	US-08-905-223-129
22	260.8	24.8	656	3	US-09-328-111-791
23	195.2	18.5	449	4	US-09-270-767-14670
24	169.2	16.1	366	3	US-08-307-499-13
25	169.2	16.1	366	3	US-09-299-268-13
26	141.4	13.4	373	4	US-09-401-064-337
27	139.8	13.3	373	4	US-09-401-064-319

C	28	115.2	10.9	366	4	US-09-401-064-307	Sequence 307, App
	29	112.4	10.7	297	4	US-09-313-294A-4435	Sequence 4435, Ap
	30	93.8	8.9	11820	4	US-09-949-016-11887	Sequence 11887, A
	31	93.8	8.9	11826	4	US-09-949-016-13767	Sequence 13767, A
	32	92	8.7	601	4	US-09-949-016-20615	Sequence 20615, A
	33	92	8.7	601	4	US-09-949-016-69028	Sequence 69028, A
	34	76	7.2	276	4	US-09-313-294A-3437	Sequence 3437, Ap
	35	71.8	6.8	304	4	US-09-313-294A-7143	Sequence 7143, Ap
	36	68	6.5	276	4	US-09-313-294A-4638	Sequence 4638, Ap
	37	61	5.8	1053	4	US-09-902-540-2666	Sequence 2666, Ap
C	38	61	5.8	13706	4	US-09-902-540-1124	Sequence 1124, Ap
	39	60.8	5.8	1141	4	US-09-806-708B-22	Sequence 22, Appl
	40	60	5.7	601	4	US-09-949-016-69031	Sequence 69031, A
	41	54.4	5.2	288	4	US-09-313-294A-6108	Sequence 6108, Ap
	42	53.8	5.1	276	4	US-09-313-294A-3244	Sequence 3244, Ap
	43	50	4.7	1230025	4	US-09-198-452A-1	Sequence 1, Appli
	44	50	4.7	1230230	4	US-09-438-185A-1	Sequence 1, Appli
C	45	47	4.5	124884	4	US-09-661-596A-76	Sequence 76, Appl

ALIGNMENTS

RESULT 1

US-10-019-733-12
; Sequence 12, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-12

Query Match	100.0%;	Score 1053;	DB 4;	Length 1053;
Best Local Similarity	100.0%;	Pred. No. 6.8e-309;	Mismatches 0;	Indels 0;
Matches 1053;	Conservative 0;			
Qy	1	ATGGGGCACC	CGGAAAGCGCGGCTGGATCAGGATCAGGATCATCTTCA	60
Db	1	ATGGGGCACC	CGGAAAGCGCGGCTGGATCAGGATCAGGATCATCTTCA	60
Qy	61	GACACCAACGAAAGTGAATTAAGTCAATTAAGACGACCTCTCTAAGAAAGAGTTCTCGC	120	
Db	61	GACACCAACGAAAGTGAATTAAGTCAATTAAGACGACCTCTCTAAGAAAGAGTTCTCGC	120	
Qy	121	CGTTTGTCTATCTTCCATCCAGTACCTGTATTTGGAAATGTATAACAGGCACAG	180	
Db	121	CGTTTGTCTATCTTCCATCCAGTACCTGTATTTGGAAATGTATAACAGGCACAG	180	
Qy	181	GCTTCTCTTGGACAGCAGAAAGTTGACTTATCAAGAGTCTCCCTCAGTGGAAACAG	240	
Db	181	GCTTCTCTTGGACAGCAGAAAGTTGACTTATCAAGAGTCTCCCTCAGTGGAAACAG	240	
Qy	241	CTTAAAGCAGATGAGAAGTACTTCTCATCTCTACATCTTTAGCCTTTTTCAGCCAGTGAT	300	
Db	241	CTTAAAGCAGATGAGAAGTACTTCTCATCTCTACATCTTTAGCCTTTTTCAGCCAGTGAT	300	
Qy	301	GGAAATGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGGTTCCAGAGGCT	360	

Db 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
QY 361 CGCTGTTCTATGGCTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTCTATGGCTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAAATTTTAAATGCAATTTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAAATTTTAAATGCAATTTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTCGATGAGATGAGATAGAAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTCGATGAGATGAGATAGAAAA 540
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTGTAAGAGGAGTTTCTCTCAGGA 600
Db 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTGTAAGAGGAGTTTCTCTCAGGA 600
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC 660
Db 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC 660
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
QY 721 TACTTAGTAAATAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 721 TACTTAGTAAATAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 780
QY 781 ATTGAGCAGAGTCTTTTAAAGAGGCTTGCAGATTGGCTCATTCGGAATGAATGCAAT 840
Db 781 ATTGAGCAGAGTCTTTTAAAGAGGCTTGCAGATTGGCTCATTCGGAATGAATGCAAT 840
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGTGAACTTGGATTCTCA 900
Db 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGTGAACTTGGATTCTCA 900
QY 901 AAGGTTTTTCAGCAGAAAAATCCTTTGATTTTATGAGAAAAATTCCTTTAGAGAAAA 960
Db 901 AAGGTTTTTCAGCAGAAAAATCCTTTGATTTTATGAGAAAAATTCCTTTAGAGAAAA 960
QY 961 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1020

RESULT 2
US-10-019-733-4
; Sequence 4, Application US/10019733
; Patent No. 682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019, 733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens

; FEATURE:
; OTHER INFORMATION:
US-10-019-733-4
Query Match 100.0%; Score 1053; DB 4; Length 1081;
Best Local Similarity 100.0%; Pred. No. 6.9e-309;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGGCGCCCGGAAAGCGGCGGCGGCTGGATCAGATGAGATCATCTTCA 60
Db 20 ATGGGGCGCCCGGAAAGCGGCGGCGGCTGGATCAGATGAGATCATCTTCA 79
QY 61 GACCAACAGAAAGTGAATAAAGTCAAAATGAAGAGCACTCTCTAAGAAAGAGTTCTCGC 120
Db 80 GACCAACAGAAAGTGAATAAAGTCAAAATGAAGAGCACTCTCTAAGAAAGAGTTCTCGC 139
QY 121 CGGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATAAACAGGCAAG 180
Db 140 CGGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATAAACAGGCAAG 199
QY 181 GCTTCCTCTGACACAGCAGAAAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
Db 200 GCTTCCTCTGACACAGCAGAAAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 259
QY 241 CTTAAAGCAGATGAGAAAGTACTTCACTCTCACAATCTTAGCCTTTTTCAGCCAGTAT 300
Db 260 CTTAAAGCAGATGAGAAAGTACTTCACTCTCACAATCTTAGCCTTTTTCAGCCAGTAT 319
QY 301 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAAGTTCAGAGGCT 360
Db 320 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAAGTTCAGAGGCT 379
QY 361 CGCTGTTCTATGGCTTTCAAAATCTCATCGAAGTGTTCACCTCAGAGATGTCAGATTG 420
Db 380 CGCTGTTCTATGGCTTTCAAAATCTCATCGAAGTGTTCACCTCAGAGATGTCAGATTG 439
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTAATGCAATTTGAA 480
Db 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTAATGCAATTTGAA 499
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTGCATGAGATGAGATAGAAAA 540
Db 500 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTGCATGAGATGAGATAGAAAA 559
QY 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGTGTAAGAGGAGTTTCTTCTCAGGA 600
Db 560 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGTGTAAGAGGAGTTTCTTCTCAGGA 619
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACCTTTTCC 660
Db 620 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACCTTTTCC 679
QY 661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTCCCTGATGTTCCAA 720
Db 680 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTCCCTGATGTTCCAA 739
QY 721 TACTTAGTAAATAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 740 TACTTAGTAAATAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 799
QY 781 ATTGAGCAGAGTCTTTTAAAGAGGCTTGCAGATTGGCTCATTTGGAATGAATGCAAT 840
Db 800 ATTGAGCAGAGTCTTTTAAAGAGGCTTGCAGATTGGCTCATTTGGAATGAATGCAAT 859
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGTGAACTTGGATTCTCA 900
Db 860 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGTGAACTTGGATTCTCA 919
QY 901 AAGGTTTTTCAGCAGAAAAATCCTTTGATTTTATGAGAAAAATTCCTTTAGAGAAAA 960
Db 920 AAGGTTTTTCAGCAGAAAAATCCTTTGATTTTATGAGAAAAATTCCTTTAGAGAAAA 979
QY 961 ACAAAATTTCTTTGAGAAACAGATTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1020

Db 980 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGAGTATGCGAGAAACC 1039
Qy 1021 ACAGATAACGCTCTTACACCTTTGGATGCGAGATTTT 1053
Db 1040 ACAGATAACGCTCTTACACCTTTGGATGCGAGATTTT 1072

RESULT 3
US-10-019-733-2
; Sequence 2, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-2

Query Match 99.8%; Score 1051.4; DB 4; Length 1053;
Best Local Similarity 99.9%; Pred. No. 2.1e-308;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGCGGACCCGGAAGGCGGAGCGCGCGGCTGGATCAGAGTATGCGAGATCATCTTCA 60
Db 1 ATGGCGGACCCGGAAGGCGGAGCGCGCGGCTGGATCAGAGTATGCGAGATCATCTTCA 60
Qy 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Qy 121 CGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG 180
Db 121 CGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG 180
Qy 181 GCTTCTCTCTGACACAGAGAGGTTGACTTATCAAGAGTCTCCCTCACTGGACACAG 240
Db 181 GCTTCTCTCTGACACAGAGAGGTTGACTTATCAAGAGTCTCCCTCACTGGACACAG 240
Qy 241 CTTAAGCAGATGAGAAGTACTTCTCTCATCTCTAGCGCTTTTTCAGCGAGTGTAT 300
Db 241 CTTAAGCAGATGAGAAGTACTTCTCTCATCTCTAGCGCTTTTTCAGCGAGTGTAT 300
Qy 301 GGAATTTGAATGAAAATTTGGTGAGCGCTTTAGTCAGGAGGTCCAGGTTCCAGAGGCT 360
Db 301 GGAATTTGAATGAAAATTTGGTGAGCGCTTTAGTCAGGAGGTCCAGGTTCCAGAGGCT 360
Qy 361 CGCTGTTCTATGCTTTCAAAATTTCTCATCGAAGTGTCTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTCTATGCTTTCAAAATTTCTCATCGAAGTGTCTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Qy 481 ACCATGCCCTATGTTTAAAGAAAAAGCAGATTTGGGCTTTGCCATGGATAGCAGATAGAAA 540
Db 481 ACCATGCCCTATGTTTAAAGAAAAAGCAGATTTGGGCTTTGCCATGGATAGCAGATAGAAA 540

Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCCCTTTGCTGCTGTAGAGAGAGTTTCTTCTCAGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCCCTTTGCTGCTGTAGAGAGAGTTTCTTCTCAGA 600
Qy 601 TCTTTTGTCTCTATATCTGGCTAAAGAGAGGTTCTTATGCCAGGACTCACATTTTCC 660
Db 601 TCTTTTGTCTCTATATCTGGCTAAAGAGAGGTTCTTATGCCAGGACTCACATTTTCC 660
Qy 661 AATGAACTCATCAGCAGAGATGAAGGACTTCTACTGTGACTTTTGTGCTGCTGATGTTCCAA 720
Db 661 AATGAACTCATCAGCAGAGATGAAGGACTTCTACTGTGACTTTTGTGCTGCTGATGTTCCAA 720
Qy 721 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGCTCAGGAGAGTCAATTGTTGATGCTGTCAA 780
Db 721 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGCTCAGGAGAGTCAATTGTTGATGCTGTCAA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAACCTTCCAGAGTTGGCCCTCATTTGGAATGAATTGCATT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAACCTTCCAGAGTTGGCCCTCATTTGGAATGAATTGCATT 840
Qy 841 TTGATGAACAGTACATTTAGCTGAGAGTCTTGTGGAACCTTGGATTTCTCA 900
Db 841 TTGATGAACAGTACATTTAGCTGAGAGTCTTGTGGAACCTTGGATTTCTCA 900
Qy 901 AAGTTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGAAAACATTTCTTTAGAGAGAAA 960
Db 901 AAGTTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGAAAACATTTCTTTAGAGAGAAA 960
Qy 961 ACAAAATTTCTTTCAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGAGTATGCGAGAAACC 1020
Db 961 ACAAAATTTCTTTCAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGAGTATGCGAGAAACC 1020
Qy 1021 ACAGATAACGCTCTTCACTTGGATGCGAGATTTT 1053
Db 1021 ACAGATAACGCTCTTCACTTGGATGCGAGATTTT 1053

RESULT 4
US-10-019-733-3
; Sequence 3, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 3
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-3

Query Match 99.8%; Score 1051.4; DB 4; Length 4955;
Best Local Similarity 99.9%; Pred. No. 5.1e-308;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGCGGACCCGGAAGGCGGAGCGCGCGGCTGGATCAGAGTATGCGAGATCATCTTCA 60
Db 245 ATGGCGGACCCGGAAGGCGGAGCGCGCGGCTGGATCAGAGTATGCGAGATCATCTTCA 304
Qy 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Db 305 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 364

121 CGTTTGTTCATCTTCCCAATCCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG 180
Db
365 CGTTTGTTCATCTTCCCAATCCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG 424
Qy 181 GCTTCCTTCTGACAGCAGAGAGAGGTGATCTATCAAGGATCTCCCTCACTGGAAACAG 240
Db 425 GCTTCCTTCTGACAGCAGAGAGAGGTGATCTATCAAGGATCTCCCTCACTGGAAACAG 484
Qy 241 CTTAAGCAGATGAGAAAGTACTCTCATCTCTCATCTCTAGCCCTTTTTCGACCCAGTGT 300
Db 485 CTTAAGCAGATGAGAAAGTACTCTCTCATCTCTAGCCCTTTTTCGACCCAGTGT 544
Qy 301 GGAATTTGAAATGAAATTTTGTGTGAGCGCTTTAGTCAGAGGTGTCAGGTTCAGAGGCT 360
Db 545 GGAATTTGAAATGAAATTTTGTGTGAGCGCTTTAGTCAGAGGTGTCAGGTTCAGAGGCT 604
Qy 361 CGCTGTTTCTATGGCTTTCAATCTCATCGAATGTTCTACTCAGAGATGTCACAGTTTG 420
Db 605 CGCTGTTTCTATGGCTTTCAATCTCATCGAATGTTCTACTCAGAGATGTCACAGTTTG 664
Qy 421 CTGATAGACATTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 665 CTGATAGACATTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 724
Qy 481 ACCATGCCCTATGTTAAGAAAGAAAGCAGATTCGGCTTCGATGATGATGATGATGAAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAGAAAGCAGATTCGGCTTCGATGATGATGATGATGAAAA 784
Qy 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTAGAGAGGATTTTCTTCACAGA 600
Db 785 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTAGAGAGGATTTTCTTCACAGA 844
Qy 601 TCTTTTGTGCTATATCTGGCTAAAGAAAGAGAGTCTTATGCCAGAGCTACTTTTTC 660
Db 845 TCTTTTGTGCTATATCTGGCTAAAGAAAGAGAGTCTTATGCCAGAGCTACTTTTTC 904
Qy 661 AATGAACATCAGCAGAGATCAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 905 AATGAACATCAGCAGAGATCAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
Qy 721 TACTTAGTAAATAAGCCTTCAAGAAAGGCTCAGGAGATCATTTGATGCTGTCAAA 780
Db 965 TACTTAGTAAATAAGCCTTCAAGAAAGGCTCAGGAGATCATTTGATGCTGTCAAA 1024
Qy 781 ATGAGCAGAGTATTTTAAAGAGCCTTGCAGTTGGCTCATTTGGAATGATTCAT 840
Db 1025 ATGAGCAGAGTATTTTAAAGAGCCTTGCAGTTGGCTCATTTGGAATGATTCAT 1084
Qy 841 TTGATGAACACATACATTGAGTTTGTAGCTGACAGATTTACTTGTGGAATTTGATCTCA 900
Db 1085 TTGATGAACACATACATTGAGTTTGTAGCTGACAGATTTACTTGTGGAATTTGATCTCA 1144
Qy 901 AAGGTTTTTCAGGAGAGAAATCCTTTGATTTTATGGAACAATTTCTTGAAGGAAAA 960
Db 1145 AAGGTTTTTCAGGAGAGAAATCCTTTGATTTTATGGAACAATTTCTTGAAGGAAAA 1204
Qy 961 ACAAAATTTCTTGAAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGATTTAGCAGAAACC 1020
Db 1205 ACAAAATTTCTTGAAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGATTTAGCAGAAACC 1264
Qy 1021 ACAGATAAGCTCTTACCTTGGATGAGATTTT 1053
Db 1265 ACAGATAAGCTCTTACCTTGGATGAGATTTT 1297

RESULT 5

US-09-962-665-9
; Sequence 9, Application US/09962665
; Patent No. 6537759
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: POLYLIPOLYGLUTAMATE SYNTHETASE GENE SEQUENCE

; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING THE
; TITLE OF INVENTION: TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015004
; CURRENT APPLICATION NUMBER: US/09/962,665
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128, 1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc_feature
; LOCATION: 1636, 1738, 2259
; OTHER INFORMATION: n = c or t
; US-09-962-665-9

Query Match 55.1%; Score 580.6; DB 4; Length 2500;
Best Local Similarity 75.1%; Pred. No. 2.2e-165;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
Qy 89 ATGAGAGCCATCTTAAGAAGAGTTCTCGCGGTTTGTCACTTTTCCAAATCCAGTACC 148
Db 397 AGATAGCGCTGTGAGAGAAACCCCGCGCTTTGTCACTTTCCCAATCCAGTACC 456
Qy 149 CTGATATTTGGAAATGTATAACAGGCACAGGCTTCTTCTGTGACAGCAGAGGTTG 208
Db 457 ATGATATCTGGCAGATGTATAGAAGGCAGAGGCTTCTTTGGACCGCGAGGTTG 516
Qy 209 ACTTATCAAGAGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAGTCTCATCT 268
Db 517 ACCTCTCAAGGACATTCAGACCTGGGAATCCCGAAACCCGAGGAGAGATATTTTATAT 576
Qy 269 CTCACATCTAGCCTTTTTCGAGCCAGTGATGGAATGTAATGAAATTTGGTGGAGC 328
Db 577 CCATGTTCTGGCTTTCTTTTGCAGCAAGCGATGECATAGTAAATGAAATTTGGTGGAGC 636
Qy 329 GCTTTAGTCAGAGAGTGCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTTCAAAATTTCTCA 388
Db 637 GATTTAGCCAGAGATTCAGATTCAGAGAGCCGCTGTTTCTATGGCTTTCCAAATTTGCCA 696
Qy 389 TCAGAGATGTTCACTCAGAGATGTACAGTTTGTGTAGTACACTTATCATCAGAGATCCCA 448
Db 697 TGGAAACATACATTCTGAAATGTATAGTCTTCTTATTGACACTTACATAAAGATCCCA 756
Qy 449 AGAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCCCTATGTTAAGAAAAAGCAG 508
Db 757 AAGAAAGGGAATTTTCTTTCAATGCCATTTGAAACGATGCCCTTGTGTCAAGAAAGCAG 816
Qy 509 ATTGGGCTTGGATGATGATGAGATGAGAAATCTACTTTTGGGAAAGAGTGTGGCT 568

```
Db 817 ACTGGGCTTGGCTGGATGGGACAAAGAGGCTACTATGGTGAACGTGTGTAGCCT 876
QY 569 TTGCTCTGTAGAAGAGTTCCTCTCAGGATCTTTTGGCTGTATATCTGTGCTAAGA 628
Db 877 TTGCTGCACTGAGGCAATTCCTTTCCGGTCTTTTGGCTGATATCTGGCTCAAGA 936
QY 629 AGAGAGCTTATGCGCAGGCTCCTTTTCCAAATGAACCTCATCAGCAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCTGCGCTCACATTTCTAATGAACCTTATAGCAGATGAGGTT 996
QY 689 TTCACCTGTGACTTGTGCTGCTGATGTTCCAAATACCTTAGTAATAAGCCTTCAGAGAA 748
Db 997 TACACTGTGATTTGTCTGCTGATGTTCAAAACACCTGGGTACAAACCAATCGGAGGA 1056
QY 749 GGTCAGGAGGATCATTTGTGATGCTGCTCAAAATTCAGCAGGAGTTTAAACAGAGCCT 808
Db 1057 GAGTAGAGAAATAATATCAATGCTGTTCCGGATAGAACAGGAGTTCCTCAGTGGCCT 1116
QY 809 TGCAGTGTGGCCTCATTTGGAATGAATTCGCAATTTTGAATGAAACAGTACATTCAGTTGTAG 868
Db 1117 TGCTGTGAAGCTCATTTGGATGAAATTCGACTCTAATGAAGCAATACATTCAGTTGTGG 1176
QY 869 CTGACAGATTAATGTGGAACCTTGGATCTCAAAGGTTTTCAGGAGGAAATCCCTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACCTGGGTTTGTAGCAAGGTTTTCAGAGTAGAGAACCCATTG 1236
QY 929 ATTTATGAAACATTTCTTTAGAGGAAACCAATTCGATGCTTTCAGTGAACGATTTTCAG 988
Db 1237 ACTTTATGAGAAATATTTCACTGGAAGGAAAGACTAACTTCTTTGAGAGAGATGAGCG 1296
QY 989 AGTATCAGGCTTTTGCAATGATGGAGGAAACCAACAGATTAACGCTTTCACCTGGATGCG 1048
Db 1297 AGTATCAGGATGGAGTGTCAAGTCCAAACAGAGAAATCTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360
```

RESULT 6

```
US-09-963-333-9
; Sequence 9, Application US/09963333
; Patent No. 6664062
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE SYNTHASE GENE SEQUENCE VARIANCES
; TITLE OF INVENTION: HAVING UTILITY IN DETERMINING THE TREATMENT
; TITLE OF INVENTION: OF DISEASE
; FILE REFERENCE: 11926-015002
; CURRENT APPLICATION NUMBER: US/09/963,333
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128, 1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
```

```
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc_feature
; LOCATION: 1636, 1738, 2259
; OTHER INFORMATION: n = c or t
US-09-963-333-9
```

```
Query Match 55.1%; Score 580.6; DB 4; Length 2500;
Best Local Similarity 75.1%; Pred. No. 2.2e-165;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
QY 89 ATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGCCGGTTGTGCATCTTTTCCAATCCAGTACC 148
Db 397 AGGATGAGCCGCTGCTGAGAGAAAACCCCGCCGCTTTGTCTATCTTCCCATCGAGTACC 456
QY 149 CTGATATTTGAAAATGTATTAACAGGACAGGCTTCTCTTGACAGCAGAGAGGTTG 208
Db 457 ATGATATCTGSCAGATGTATAAGAAAGCAGAGGCTTCTCTTTGGACCCCGAGGAGTTG 516
QY 209 ACTTATCAAGAGTCTCCCTCACTGGAACAAGCTTTAAAGCAGATGAGAAGTACTTTCATCT 268
Db 517 ACCTCTCNAAGGACATTCAGCACTGGGAATCCCTCGAACCAGGAGAGATATTTTATAT 576
QY 269 CTCACATCTTAGCCCTTTTTCAGCCAGTGAATGTAATGTAATGTAATGTAATGTAATGTA 328
Db 577 CCATGTTCTGGCTTCTTTTTCAGCAGCAGATGCAATGTAATGTAATGTAATGTAATGTA 636
QY 329 GCTTTAGTCAGGAGGTCAGGTTTCAGAGGCTCGCTGTTCTATGCTGCTTCAAAATCTCA 388
Db 637 GATTAGCCAAGAGTTTCAATTAACAAGCCGCTGTTCTATGGCTTCCAAATTTGCA 696
QY 389 TCGAGATGTTCACTCAGAGATGTACAGTGTGCTGATAGACACTTACATCAGAGATCCCA 448
Db 697 TGAAGAACATACATCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 756
QY 449 AGAAAGGGAAATTTTATTTAATGCAATGAAACCATGCCCCTATGTTAAGAAAAAGCAG 508
Db 757 AGAAGAGGAAATTTCTCTCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 816
QY 509 ATGGGCTTTCGATGATGATAGAGATAAATCTACTTTTGGGAAAGAGTGGGCT 568
Db 817 ACTGGGCTTTCGCTGATTTGGGACAAAGAGGCTACTATGCTGTAACCTGTTGAGCCT 876
QY 569 TTGCTGCTGTAGAAGGAGTTTCTTCTCAGATGCTTTTCTGCTATATTTCTGGCTAAGA 628
Db 877 TTGCTGAGTGGAGGCAATTTCTTTTCCGGTTCCTTTTGGCTGATATTTCTGGCTCAAGA 936
QY 629 AGAGAGGCTTATGTCAGGAGCTCATCTTTTTCCAATGAACTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCTGGCTCACATTTTCTAATGAACCTTATAGCAGATGAGGTT 996
QY 689 TTCACCTGTGACTTGTGCTGCTGATGTTCCAATCTTATAGTAAATPAAGCCTTCAGAGAAA 748
Db 997 TACACTGTGATTTTGTCTGCTGATGTTCAAAACACCTGGGTACACAAACCAATCGGAGGA 1056
QY 749 GGGTCAGGAGATCATTTGTGATGCTGTCAAAATTCAGCAGGAGTGTTTTAAACAGAGCCT 808
Db 1057 GAGTAGAGAAATAATATCAATGCTGTTCCGGATAGAACAGGAGTTCCTCAGTGGCCT 1116
QY 809 TGCAGTGTGGCCTCATTTGGAATGAATTCGATGAAACAGTACATTCAGTGTGTAG 868
Db 1117 TGCTGTGAAGCTCATTTGGATGAAATTCGACTCTAATGAAGCAATACATTCAGTGTGTGG 1176
QY 869 CTGACAGATTTACTTGTGGAACCTTGGATCTTCAAGGTTTTCAGGAGGAAATCCCTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACCTGGGTTTGTAGCAAGGTTTTCAGAGTAGAGAACCCATTG 1236
QY 929 ATTTATGGAACATTTCTTTAGAGGAAAAACAAATTTCTTTGAGAAACGAGTTCAG 988
```

Db 1237 ACTTTATGAGAATATTTCTACTGGAAGGAAAGACTAACTTCTTTGAGAAGAGTAGGCG 1296
QY 989 AGTATCAGCGTTTTCAGTATGCGAAGAAACACACAGATACGCTCTTCCCTTGGATGCAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGATGTCAGTCCAAACAGAGAATTCCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 7

US-09-962-677-9
; Sequence 9, Application US/09962677
; Patent No. 6759200
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE PHOSPHORYLASE GENE SEQUENCE
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING
; TITLE OF INVENTION: THE TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015003
; CURRENT APPLICATION NUMBER: US/09/962,677
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128..1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc_feature
; LOCATION: 1636..1738, 2259
; OTHER INFORMATION: n = c or t
US-09-962-677-9

Query Match 55.1%; Score 580.6; DB 4; Length 2500;
Best Local Similarity 75.1%; Pred. No. 2.2e-165;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
QY 89 ATGAGAGGCATCTCTTAAGAAGAGTTCTGCGCGTTTGTCTATCTTTTCAATCCAGTACC 148
Db 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTGTCTATCTTCCCATCGAGTACC 456
QY 149 CTGATATTTGGAATGTATAACAGGCACAGCTTCTCTTGGACAGAGAGGTTG 208
Db 457 ATGATATCTGGCAGATGTATAAGAGGCAGAGGCTTCTTTTGGACCGCGGAGGTTG 516
QY 209 ACTTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTCACT 268
Db 517 ACCTCTCNAGGACATTGAGCACTGGGATCCCTGAAACCGGAGGAGATATTTATAT 576

QY 269 CTCACATCTTAGCCTTTTGGCAGCCAGTGCATGGAATTGTAAATGAAAAATTTGGTGGAGC 328
Db 577 CCATGTTCTGGCTTTCTTTGCGAAGCGATGCGATAGTAATGAAAAATTTGGTGGAGC 636
QY 329 GCTTTAGTCAGGAGGTGCGAGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAAATTTCTCA 388
Db 637 GATTTAGCCAAGAAGTTTCAGATTCAGAGAGCCCGCTGTTTCTATGGCTTCCAAATTTGCCA 696
QY 389 TCGAGAATGTTCACTCAGAGATGTACAGTTTGTCTGTAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAAAACATACATTCCTGAAATGTATAGTCTTCTTATTCACCTACATATAAAGATCCCA 756
QY 449 AGAAAAAGGGAATTTTATTTAATGCAATTGAAACCATGCGCTTATGTTAAGAAAAAAGCAG 508
Db 757 AGAAAAAGGGAATTTCTCTCAATGCGCATTTGAAAGATGCGCTTGTGTCAAGAAGAGGCGAG 816
QY 509 ATTGGGCTTTGCGATGAGATAGACATAGAAAAATCTACTTTTGGGGAAGAGTGGTGGCCT 568
Db 817 ACTGGGCTTTGCGCTGGAATTTGGGGAACAAGAGGCTACCTATGTTGAACGTTGTGAGCCT 876
QY 569 TTGCTGCTGAGAAGAGTTTCTCTCAGGATCTTTTCTGCTGCTATATTTCTGGCTTAAGA 628
Db 877 TTGCTGCTGAGAAGGCAATTTCTTTTCCGCTTCTTTTGGCTCGATATTTCTGGCTCAAGA 936
QY 629 AGAGAGGCTCTTATGCCAGGACTCACCTTTTCCAAATGAACCTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCTGCGCTCACATTTTCTAATGAACTTATTAGCAGAGATGAGGGTT 996
QY 689 TTCACCTGTGACTTTTGGCTGCTGATGTTCCAAATCTTAGTAAATAAGCCTTTCAGAAGAAA 748
Db 997 TACACTGTGATTTTGGCTGCTGATGTTCAAAACACCTGGTACACAAACCATCGGAGGAGA 1056
QY 749 GGGTCAGGGAGATCATTGTTGATGCTGTCAAAATGAGCAGGAGTGTTCCTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATAATATCAATGCTGTTCCGATAGAAACAGGAGTTCTCTCACTGAGGCT 1116
QY 809 TGCAGCTTGGCTCATTTGGAATGAATTCATTTTGTGTAAGCACTACATTCAGTTTGTAG 868
Db 1117 TGCCTGTGAAGCTCATTTGGATGAATTTGCACTTAATGAAGCAATACATTCAGTTTGTGG 1176
QY 869 CTGACAGATTTACTTTGTGGAACCTTTGGAATTTCTCAAGGTTTTTTCAGGAGAAAAATCCTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACCTGGGTTTGTAGCAAGTTTTCAGAGTAGAGAACCAATTTG 1236
QY 929 ATTTTATGGAACAATTTCTTTAGAGGAAAAACAATTTCTTTGAGAAAAAGAGTTTCAG 988
Db 1237 ACTTTATGGAGAATAATTCACCTGGAAGGAAAGACTAACTTCTTTGAGAAGAGATGAGCG 1296
QY 989 AGTATCAGCGTTTTCAGTTATGCGAGAAACACAGATACGTTCTTACCTTGGATGCAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGATGTCAGTCCAAACAGAGAATTCCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 8

US-09-023-655-1370
; Sequence 1370, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO

STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,655
FILING DATE: HERWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1370:
SEQUENCE CHARACTERISTICS:
LENGTH: 2500 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: G36154
US-09-023-655-1370

Query Match 55.1%; Score 580; DB 4; Length 2500;
Best Local Similarity 75.1%; Pred. No. 3.4e-165;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 89 ATGAGAGCCATCTCTAAGAGAGTTCTGCGGGTTTGTGATCTTCCATTCAGTACC 148
DB 397 AGGATGAGCGCTGCTGAGAGAAAACCCCGCGCTTGTCTCTTCCATTCAGTACC 456
QY 149 CTGATATTTGGAATAATGTAACAGCACAGGCTTCTTCTTGACAGAGAGGTTG 208
DB 457 ATGATATCTGGCAGATGTAAGAGGAGGCTTCTTCTTGACCGCGGAGGTTG 516
QY 209 ACTTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAGTCTCA 268
DB 517 ACTCTCCAAGGACATTCAGCACTGGGAATCCTGAAACCCGAGGAGAGATATTTAT 576
QY 269 CTCACATCTTAGCCTTTTGGAGCAGTGATGGAATTTGTAATGAAATTTGGTGGAG 328
DB 577 CCCATGTTCTGGCTTTCTTGGCAGCAAGCGATGGCATAGTAATGAAACTTGGTGGAG 636
QY 329 GCTTTAGTCAGAGTGCAAGTTCCAGAGCTCGCTGTTTCTATGCTTCAATCTCA 388
DB 637 GATTTAGCCAAAGTTCAGATACAGAGCGCGCTGTTTCTATGCTTCAATGCGCA 696
QY 389 TCGAGATGTTCACTCAGAGATGTACAGTTTGTGATAGACACTTACATCAGAGATCCA 448
DB 697 TGGAAACATACATCTGAAATGTAGTCTTCTTATGACACTTACATTAAGATCCA 756
QY 449 AGAAAGGGAATTTTATTAATGCAATTTGAACCAATGCTTATGTAAGAAAAAGCAG 508
DB 757 AAGAAAGGGAATTTCTTCAATGCAATTTGAACCAATGCTTATGTAAGAAAAAGCAG 816
QY 509 ATTGGGCTTGGATGATAGCAGATAGAAATCTACTTTTGGGAAAGTGGTGGCT 568
DB 817 ACTGGGCTTGGGCTTGGATTTGGGAAACAAAGAGGCTTACCTATGTTGAGCT 876
QY 569 TTGCTGCTGTAAGAGGAGTTTCTTCTCAGGATCTTTTCTGCTATATTTGCTAAAGA 628
DB 877 TTGCTGCTGTAAGAGGAGTTTCTTCTCAGGATCTTTTCTGCTGATATTTCTGGCTCA 936

QY 629 AGAGAGGTCTTATGCGCAGGACTCACTTTTTCATTAAGTCACTCATCAGCAGAGATGAAGAC 688
DB 937 AACGAGGACTGATGCGCTGCTCACAATTTCTAATGAATTTATAGCAGAGATGAGGTT 996
QY 689 TTCACGTGTGATTTGCTTGTGCTGATGTTCCAAATCTTAGTAAATAAGCCTTCAGAAAGAA 748
DB 997 TACACTGTGATTTGCTTGTGCTGATGTTCAAAACACCTGGTACACAAACCATCGGAGAGA 1056
QY 749 GGGTCAGGAGATCATTGTTGATGCTCTCAAAATTTGACAGAGCTTTTAAACAGAGCCT 808
DB 1057 GAGTAAGAGAAATAAATTAATCAATGCTGTTCGATAGAAACAGGAGTTCTCACTGAGGCT 1116
QY 809 TGCAGGTTGCGCTCATTGGAATGAATTCATTTTATGAAACAGTACATTTGAGTTTGTAG 868
DB 1117 TGCTGTGAAGCTCATTGGGATGATTCGACTCTAATGAACATATACATTTGATTTGTG 1176
QY 869 CTGACAGATTAATCTTGTGGAATCTGGAATTTCTCAAAAGGTTTTCAGGCAGAAATCTTTTG 928
DB 1177 CAGACAGACTTATGCTGGAATCTGTTTTCAGCAAGGTTTTCAGAGTAGAGAACCCATTG 1236
QY 929 ATTTTATGGAACAATTTCTTTAGAGGAAAAAATAATTTCTTTGAGAAACGAGTTTCAG 988
DB 1237 ACTTTATGGAATAATTTCTACTGGAAGGAAAGACTAACTCTTTTGAGAGAGAGTAGGCG 1296
QY 989 AGTATCAGCGCTTTTGCAGTTATGCGAGAAACACAGATAACGCTTTCACCTTTGGATGCG 1048
DB 1297 AGTATCAGAGGATGGGAGTGTCAAGTCCAAACAGAGAAATCTTTTACCTTTGATGCTG 1356
QY 1049 ATTT 1052
DB 1357 ACTT 1360

RESULT 9

US-09-949-016-145
Sequence 145, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 145
LENGTH: 2500
TYPE: DNA
ORGANISM: Human
US-09-949-016-145

Query Match 55.1%; Score 580; DB 4; Length 2500;
Best Local Similarity 75.1%; Pred. No. 3.4e-165;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 89 ATGAGAGCCATCTCTAAGAGAGTTCTGCGCGGTTTGTCTCTTCCATTCAGTACC 148
DB 397 AGGATGAGCGCTGCTGAGAGAAAACCCCGCGCTTGTCTCTTCCATTCAGTACC 456
QY 149 CTGATATTTGGAATAATGTAACAGCACAGGCTTCTTCTTGACAGAGAGGTTG 208
DB 457 ATGATATCTGGCAGATGTAAGAGGAGGCTTCTTCTTGACCGCGGAGGTTG 516
QY 209 ACTTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAGTCTCACT 268

Db 517 ACCTCTCCAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
QY 269 CTCACATCTTAGCCTTTTTCAGAGCCAGTGAGTAATGTAATGAAATTTGGTGAGC 328
Db 577 CCATGTTCTGGCTTTCTTTCAGCAAGGATGCGCATAGTAATGAAACTTTGGTGAGC 636
QY 329 GCTTTAGTCAGGAGTGCGAGTTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTCCTCA 388
Db 637 GATTTAGCAAGAAAGTTTCAGATTACAGAAAGCCGCTGTTTCTATGCTTCCAAATTCGCA 696
QY 389 TCAGAAATGTTCACTCAGAGATGTAAGTTTTCAGATGTAAGATTCATCAGATGTAAGATTC 448
Db 697 TCGAAATACATACATCTGAAATGTAAGTTTTCAGATGTAAGATTCATCAGATGTAAGATTC 756
QY 449 AGAAAGGAAATTTTATTTAATGCAATTTGAACCAATGCAATGCAATGCAATGCAATGCAAT 508
Db 757 AGAAAGGAAATTTTCTTCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 816
QY 509 ATTGGGCTTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 568
Db 817 ACTGGGCTTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 876
QY 569 TTGCTGCTGATAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 628
Db 877 TCGCTGCTGATAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 936
QY 929 ATTTTATGAAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
Db 1237 ACTTTATGAGAATAATTTCACTGGAAGGAAAGACTTAATTTCTTTGAGAAAGAGTAGGCG 1296
QY 989 AGTATCAGCGTTTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1048
Db 1297 AGTATCAGAGGATGGGAGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 10

US-09-949-016-2025
; Sequence 2025, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2025
; LENGTH: 2479
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-2025

Query Match 54.9%; Score 578.4; DB 4; Length 2479;
Best Local Similarity 75.0%; Pred. No. 1e-164;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 89 ATCAAGAGCCACTCTTAAGAAAGATCTCCGCGGTTTGTCTCATCTTTCCATCCAGTACC 148
Db 397 AGGATGAGCGGCTGCTGAGAGAAAACCCCGCGCTTTGTCTCATCTTCCCATCCAGTACC 456
QY 149 CTGATATTTGGAAAATGTATAAACAGGCACAGGCTTCTTCTCGACAGCAAGAGGTTG 208
Db 457 ATGATATCTGGCAGATGTAAGAGGACAGGCTTCTTTTGGACCGCCGAGAGGTTG 516
QY 209 ACTTATCAAGAGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTACTTCACT 268
Db 517 ACCTCTCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
QY 269 CTCATATCTTAGGCTTTTTCAGCCAGTGAATGTAATGAAATTTGGTGAGC 328
Db 577 CCATGTTCTGGCTTTCTTTGACAGCAAGCGATGCGCATAGTAAATGAAACCTTGGTGAGC 636
QY 329 GCTTTAGTCAGAGGTCAGGTTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTCCTCA 388
Db 637 GATTTAGCAAGAAAGTTTCAGATTACAGAACCCGCTGTTTCTATGGCTTCCAAATTCGCA 696
QY 389 TCAGAAATGTTCACTCAGAGATGATACAGATTTTGTGTATGATAGACACTTACATCAGAGATCCCA 448
Db 697 TCGAAACATACATCTGAAATGTAAGTTTCTTATGACACTTACATAAAGATCCCA 756
QY 449 AGAAAGGAAATTTTATTTAATGCAATTTGAAACCATGCGCTTATGTTAAGAAAGAGCAG 508
Db 757 AGAAAGGAAATTTCTTCTCAATGCAATTTGAAACGATGCGCTTGTGTCAAGAAAGGAGCAG 816
QY 509 ATTGGGCTTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 568
Db 817 ACTGGGCTTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 876
QY 569 TTGCTGCTGATAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 628
Db 877 TCGCTGCTGATAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 936
QY 629 AGAGAGTCTTATGCGAGGATCTCATTTTTCAAATGCAATGCAATGCAATGCAATGCAATG 688
Db 937 AACGGAATGATGCTGCGCTCACAATTTTCTAATGAACTTATTTAGCAGAGATGAGGTT 996
QY 689 TTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 748
Db 997 TACACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1056
QY 749 GGGTCAGGAGATCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 808
Db 1057 GAGTAGAGAAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1116
QY 809 TGCAGTTGGCTCATTTGGAAATGAAATTTGATGATGATGATGATGATGATGATGATGATG 868
Db 1117 TGCCTGTAAGGCTCATTTGGGATGAAATTTGCACTCTAATGAAAGCAATACATTTGAGTTG 1176
QY 869 CTGACAGATTTACTTTGGAACTTTGGATTTCTCAAGGTTTTCAGGCGAGAAATCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAATCTGGGTTTATGCAAGGTTTTCAGAGTAGAGAACCCATTTG 1236
QY 929 ATTTTATGAAACATTTCTTTTATGAGAGGAAACAAATTTCTTTTATGAGAAACGAGTTTCAG 988

Db 1237 ACTTTATGAGAATAATTTCACTGGAGGAAAGACTAACTTTCTTTGAGAGAGAGTAGGCG 1296
QY 989 AGTATCAGCGTTTGTGAGTTATGGCAGAAACACACAGATAACGCTTTCACCTTGGATGCGAG 1048
Db 1297 AGTATCAGAGGATGGAGTGAATGTAAGTCCAAACAGAGAATTTCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 11

US-08-905-223-125
; Sequence 125, Application US/08905223
; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duelt, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobb, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 125:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 481 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: CDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Brain
; FEATURE:

NAME/KEY: sig_peptide
; LOCATION: 41..343
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 4.4
; OTHER INFORMATION: seq ISHILAFPAASDG/IV
US-08-905-223-125

Query Match 41.7%; Score 439.4; DB 3; Length 481;
Best Local Similarity 99.8%; Pred. No. 5.9e-123;
Matches 440; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGGGAGCCCGAAGCGCGAAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
Db 41 ATGGGGGAGCCCGAAGCGCGAAGCGCGGCTGGATCAGATGAGATCATCTTCA 100
QY 61 GACACCAAGGAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAGAGTTCTCGC 120
Db 101 GACACCAAGGAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAGAGTTCTCGC 160

QY 121 CGGTTTGTCTATCTTTTCAATCCAGTACCTGATATTTGGAAAAATGTATAAACAGGCACAG 180
Db 161 CGGTTTGTCTATCTTTTCAATCCAGTACCTGATATTTGGAAAAATGTATAAACAGGCACAG 220
QY 181 GCTTCCCTTTGGAGCAGCAGAGAGGTTGACTTATCAAAAGGATCTCCCTCCTCCTGGAACAAG 240
Db 221 GCTTCCCTTTGGAGCAGCAGAGAGGTCGACTTATCAAAAGGATCTCCCTCCTCCTGGAACAAG 280
QY 241 CTTAAGCAGATGAGAAGTACTTCTCTCAGATCTTAGCGCTTTTTCGAGCCAGTGAT 300
Db 281 CTTAAGCAGATGAGAAGTACTTCTCTCAGATCTTAGCGCTTTTTCGAGCCAGTGAT 340
QY 301 GGAATTGTAATGAAAAATTTGGTGAGCGCTTTTAGTCAGAGAGTGCGAGGTTCCAGAGGCT 360
Db 341 GGAATTGTAATGAAAAATTTGGTGAGCGCTTTTAGTCAGAGAGTGCGAGGTTCCAGAGGCT 400
QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 401 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 460
QY 421 CTGATAGACACTTACATCAGA 441
Db 461 CTGATAGACACTTACATCAGA 481

RESULT 12

US-08-307-499-1/c
; Sequence 1, Application US/08307499
; Patent No. 5651972
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,499
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF35.1.FWCCI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800

INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:


```
; LENGTH: 14176 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3852..4226
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4585..4887
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5131..5310
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5760..5912
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6786..7130
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 10148..10513
; US-08-307-499-1
```

```
Query Match 40.8%; Score 429.8; DB 1; Length 14176;
Best Local Similarity 66.2%; Pred. No. 3.3e-119;
Matches 551; Conservative 0; Mismatches 327; Indels 5; Gaps 2;

QY 71 AAAGTGAATAAAGTCAAAATGAAGACCCACTCTCTAAGAAAGATGTCCTCCCGGTTGTGCA 130
Db 10841 AAAATGAATTTTATACACGAATGGAGCCCTATTCTTCAAGAGTCTGATTCAGGTTGTTA 10782

QY 131 TCTTTCCATCCAGTACCCTGATATTTCGAAATGTATAACAGGACACAGGCTCTCTTCT 190
Db 10781 TTTTCCCTTATTAAGTATCATGTATCTGGAATATGTATAACAAATCAGTGGCAAGTTT 10722

QY 191 GGACAGCAGCAGAGGTGACTTATCAAGAGTCTCCCTCACTGGCAACAGCTTAAAGCAG 250
Db 10721 GGACCGTTGAAGAGTAGATTTTATCAAGAGTTTAGATGTTGGGATAAATTAACCTAAG 10662

QY 251 ATGAGAGTACTTCACTCTCACTCTTAGCCCTTTTTCACGCCAGTGTAGGAATGTAA 310
Db 10661 ACGAAAAATCTTTATAAAACATATACCTAGCATTTTTTTCATCTAGTGATGTATTGTAA 10602

QY 311 ATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCTCGCTGTTCT 370
Db 10601 ATGAGAAATTTAGCGGAAAGATTTTATGTGATGTACAGTGTTCAGAGGCGACGATGTTCT 10542

QY 371 ATGGCTTTCAAAATCTCATCGAGAATGTTTCACTCAGAGATGTACAGTGTGTGATAGCA 430
Db 10541 ATGGATTTCAAAATGCTATGGAAAATTTTCACTCAGAAAATGTATAGTTTATTAAATAGATA 10482

QY 431 CTATCATCAGAGATCCCAAGAAAGGAAATTTTTTAAATCAATGTAAACCATGCCCT 490
Db 10481 CATATGTAAGAGATAATATAGAAAATGCAATTTATTTAAACGCTATAGAAAACATGGAAT 10422

QY 491 ATGTTAAGAAAAACAGATTTGGCGCTTCGGATGTAGCAGATAGAAAATCTACTTTTG 550
Db 10421 GCGTAAAAAAGAAAGCTGATTTGGCGGCAGAAAATGGATATC---TAGCAACAAGGTATATG 10365

QY 551 GGGAAAGAGTGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGATCTTTTGCTG 610
Db 10364 GAGAAGAGTAGTAGCATTTTGCAGCTGTGGAGGGAATATCTTTCTGTTCAATTTGCTG 10305

QY 611 CTATATCTGGCTAAAGAGAGAGGTCTTATGCGCAGGACTCACTTTTCCAATGAACCTCA 670
Db 10304 CTATATTTGGATAAAAAACAGGATTTGATGCCGGATACATTTTCTATGACTAA 10245

QY 671 TCAGCAGAGATGAAGCACTTCACTGTGACTTTGCTTGTGCTGATGTTCCAAATCTACTAGTAA 730
Db 10244 TAAGTAGAGACGAAGGTTTACATTTGTGATTTTGGCGTGTAAATGTATTTAAACATTTATTAC 10185
```

RESULT 13

```
US-08-307-499-14
; Sequence 14, Application US/08307499
; Patent No. 5651972
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,499
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14176 base pairs
```

```
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
```

	Query Match	40.8%	Score 429.8	DB 1	Length 14176
	Best Local Similarity	66.2%	Pred. No. 3.3e-119		
	Matches 651	Conservative 0	Mismatches 327	Indels 5	Gaps 2
QY	71	AAAGTGAATAAAGTCAATGAAGAGCCACTCTAAGAAAAGATTCTCGCGGGTTTGTCA	130		
DB	3336	AAAATGATTTTATACAGCAATGAGCGCTATTCTTCAAGAGTCTGATTCTAGGTTCTGTTA	3395		
QY	131	TCTTTCCAAATCCAGTACCCCTGATATTGGAAAAATGTATAACGACACAGGCTTCCCTCT	190		
DB	3396	TTTTCCCTATTAGTATCATGATCTCGAAAAATGTATAACAAATCATAGTGGCAAGTTTTT	3455		
QY	191	GGACAGCAGAGAGGGTTGACTTATCAAGGATCTCCCTCACTGGAAACAAGCTTAAAGCAG	250		

[illegible]

RESULT 14
US-09-299-268-1/c
; Sequence 1, Application US/09299268
; Patent No. 6217882
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eliadio
; APPLICANT: Gibbs, B.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,268
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/901,127
FILING DATE:
APPLICATION NUMBER: US 07/908,241
FILING DATE: 1-JUL-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/908,630
FILING DATE: 29-JUN-1992
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/342,212
FILING DATE: 21-APR-1992
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Salliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 14176 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 3852..4226
FEATURE:
NAME/KEY: CDS
LOCATION: 4585..4887
FEATURE:
NAME/KEY: CDS
LOCATION: 5131..5310
FEATURE:
NAME/KEY: CDS
LOCATION: 5760..5912
FEATURE:
NAME/KEY: CDS
LOCATION: 6786..7130
FEATURE:
NAME/KEY: CDS
LOCATION: 10148..10513
US-09-299-268-1

Query Match 40.8%; Score 429.8; DB 3; Length 14176;

Best Local Similarity 66.2%; Pred. No. 3.3e-119; Indels 5; Gaps 2;
Matches 651; Conservative 0; Mismatches 327;

QY 71 AAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTCTCGCGCGTTGTCA 130
DB 10841 AAATGAATTTTATACACGAAGAGCCCTATCTTCAAGAGTCTGATCTAGGTTGTTA 10782
QY 131 TCTTTCCAAATCCAGTACCTCGATATTGTGAAATATGTAACAGGACAGGCTTCCTTCT 190
DB 10781 TTTTCCCTATTAAATATCATATCTCGAAATATGTAACCAATCAGTGGCAAGTTTT 10722

QY 191 GGACAGCAGAGAGAGGTGACTTATCAAGAGATCCCTCACTGGAACAAGCTTAAAGCAG 250
DB 10721 GGACCGTTGAAGAAGTAGATTTATCAAAAGATTAGATGATGGATAAATTAACATAAG 10662
QY 251 ATGAGAAGTACTTCACTCTCACATCTTAGCCTTTTTTGAGCCAGTGTAGTAATGTAA 310
DB 10661 ACGAAAAAATACCTTTATAAAACATATACCTAGCATTTTTTGTGATCTAGTGTATGTAA 10602
QY 311 ATCAAAATTTGGTGGAGCGCTTTTAGTCAGAGGTGACAGGTTCAGAGGCTCGCTCTTCT 370
DB 10601 ATGAGAATTTAGCGAAAGATTTTATGTGATGTACAGTGTTCAGAGGACAGATGTTCT 10542
QY 371 ATGCTTTTCAAAATCTCATCGAAGTGTTCATCGAGAGATGTACAGTTTGTGTATAGACA 430
DB 10541 ATGGAATTTCAATAGCTATGAAAAATATTCATTGAGAAATGTAAGTTTTATTATAGATA 10482
QY 431 CTTACATCAGAGATCCCAAGAAAGGGAATTTTTTATTATTAATGCAATTGAAACCACTGCCCT 490
DB 10481 CATATGTAGAGATAATATAGAAAAAATGCAATTTATTTAAAGCTATAGAAAAAATGGAAT 10422
QY 491 ATGTTAAGAAAAAGCAGATTTGGGCCCTTGGCATGATAGAGATAGAAAAATCTACTTTTG 550
DB 10421 GCGTAAAAAAGAAAGCTGATTTGGGCCAGAAAAATGGATATC---TAGCAACAGGTATATG 10365
QY 551 GGAAAGAGTGTGGCCCTTTTGTCTGTAGAAAGAGTTTTTCTCTCAGAGATCTTTTGTCTG 610
DB 10364 GAGAAAGAGTAGTAGCATTTTGCAGCTGTGGAGGAATATCTTTTCTGGTTCAATTTGCTG 10305
QY 611 CTATATTTCTGGCTAAAGAGAGAGGTCTTATGCGCAGGACTCACATTTTCCAAATGAACCTCA 670
DB 10304 CTATATTTTGGATAAAAAAAGAGGATTTGATGCGCGGATTAACATTTTCTAATGAACCTAA 10245
QY 671 TCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAATCTAGTAA 730
DB 10244 TAAGTAGAGACGAAGGTTTACATTTGCGTGTGTTTAATGTTTAAACATTTTATTATAC 10185
QY 731 ATAAAGCTTCAGAAAGAGGTCAGGAGAGATCATTTGTTGATGCTGTCAAAATTCAGCAGG 790
DB 10184 ATCCACCATCTAAGGAAGTTATTAACGTCGATATCATTTGATCGGTTAATATAGAAAGG 10125
QY 791 AGTTTTTACAGAAAGCCTTCCAGTGTGGCCTCATTTGGAATGAATTTGCAATTTGATGAAC 850
DB 10124 AGTTTTTGACAGTTGCTTATTCGCGTGGATCTTATAGGTATGAATTTGTTGTTAATGTCTC 10065
QY 851 AGTACATTGAGTTGTAGCTGACAGATTTACTTTGGAACTTGGATTTCTCAAGAGTTTTTC 910
DB 10064 AGTATATAGAAATTCGTCGAGATAGATTTATTAACAGAGTTAGGTTGTGAAAG--TCTCA 10007
QY 911 AGGCAAAAAATCCTTTTGTGATTTTATGAAAAACATTTCTTTAGAAAGAAAAACAAATTTCT 970
DB 10006 ATGTATATATCTTTTGTAGCTTTTATGAGTATATATCACTAGAGGTAAAGACTAATTTT 9947
QY 971 TTGAGAAACGAGTTTCAGAGTATCAGGCTTTTTCAGTTATGGCAGAAACCAAGATAACG 1030
DB 9946 TCGAACGACGAGTTAGTGAATATCAAAAGATGGGGTGTGTTTACAAATAAGAGAGAGATA 9887
QY 1031 TCTTCACTTGGATGCAATTTT 1053
DB 9886 TATTTTCTACGGATATAGATTTT 9864

RESULT 15

US-09-299-268-14

; Sequence 14, Application US/09299268

; Patent No. 6217882

; GENERAL INFORMATION:

; APPLICANT: Moyer, Richard W.

; APPLICANT: Vi uela, Eladio

; APPLICANT: Gibbs, E.P.J.

; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a

; TITLE OF INVENTION: Live Vaccine Vector

; NUMBER OF SEQUENCES: 60

; CORRESPONDENCE ADDRESS:

ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,268
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/901,127
FILING DATE:
APPLICATION NUMBER: US 07/908,241
FILING DATE: 1-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,630
FILING DATE: 29-JUN-1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/342,212
FILING DATE: 21-APR-1992
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF35.1.FWCCI
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 14176 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 138..1460
FEATURE:
NAME/KEY: CDS
LOCATION: 2456..2659
FEATURE:
NAME/KEY: CDS
LOCATION: 2809..3030
FEATURE:
NAME/KEY: CDS
LOCATION: 3070..3330
FEATURE:
NAME/KEY: CDS
LOCATION: 3356..4180
FEATURE:
NAME/KEY: CDS
LOCATION: 4392..5894
FEATURE:
NAME/KEY: CDS
LOCATION: 6171..6398
FEATURE:
NAME/KEY: CDS
LOCATION: 6447..6875
FEATURE:
NAME/KEY: CDS
LOCATION: 6928..7431
FEATURE:
NAME/KEY: CDS
LOCATION: 7454..7858
FEATURE:

NAME/KEY: CDS
LOCATION: 7895..8155
FEATURE:
NAME/KEY: CDS
LOCATION: 8215..8682
FEATURE:
NAME/KEY: CDS
LOCATION: 8715..9539
FEATURE:
NAME/KEY: CDS
LOCATION: 9562..10272
FEATURE:
NAME/KEY: CDS
LOCATION: 10316..11908
FEATURE:
NAME/KEY: CDS
LOCATION: 11971..12780
FEATURE:
NAME/KEY: CDS
LOCATION: 12829..13107
FEATURE:
NAME/KEY: CDS
LOCATION: 13149..14171
US-09-299-268-14
Query Match 40.8%; Score 429.8; DB 3; Length 14176;
Best Local Similarity 66.2%; Pred. No. 3.3e-119; Indels 5; Gaps 2;
Matches 651; Conservative 0; Mismatches 327;
QY 71 AAAGTGAATAAAGTCAAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGCGGTTTGTCA 130
DB 3336 AAAATGATTTTATACACGAATCGAGCCTATTCTCAAGAGTCTGATCTAGGTTCTGTTA 3395
QY 131 TCTTTCCAATCCAGTACCCCTGATATTTGGAAAAATGTATAAACAGGCAAGGCTTCTCTTCT 190
DB 3396 TTTTCCTATTAAAGTATCATGATATCTGGAAAAATGTATAAACAACTCAGTGGCAAGTTTTT 3455
QY 191 GGACAGCAGAGAGGTTGACTTATCAAGAGATCCCTCACTGGAACAGCTTAAGCAG 250
DB 3456 GGACCGTTGAAGAAAGTAGATTTTATCAAAAGATTAGATGATGGGATAAATAACTAAAG 3515
QY 251 ATGAGAAGTACTTCACTCTCTCACATCTTAGCCCTTTTTCAGAGCCAGTATGGAATTGTA 310
DB 3516 ACGAAAAATCTTTATAAAACATATACTAGCATTTTTCAGATCTAGTATGATTTGTA 3575
QY 311 ATGAAAAATTTGGTGGAGCGCTTTTATGACAGAGTGACAGGTTCCAGAGGCTCGCTTTCT 370
DB 3576 ATGAGAATTTAGCGGAAAGATTTTATGTCGATGTACAGTGTTCAGAGGCAAGTGTCT 3635
QY 371 ATGGCTTTCAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTGTGATAGACA 430
DB 3636 ATGGATTTCAAATAGCTATGGAATAATTTCAATTCAGAAATGTATATTAATAGATA 3695
QY 431 CTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAACACCATGCCCT 490
DB 3696 CATATGTAAGAGATAATATAGAAAAAATGCAATTTATTAACGCTATAGAAACAATGGAAT 3755
QY 491 ATGTTAAGAAAAAGCAGATTTGGGCTTCGATGGATAGCAGATAGAAAAATCTACTTTTG 550
DB 3756 GCGTAAAAAAGAAAGCTGATTTGGCCAGAAAAATGGATATC--TAGCAACAAGGTATATG 3812
QY 551 GGGAAAGAGTGGTGGCTTTTCGCTGTGTAGAAAGAGTTTCTCTCTCAGGATCTTTTGTG 610
DB 3813 GAGAAAGAGTAGTAGCATTTTGCAGCTGTGGAGGGAATATTTCTTTCTGTTCAATTTGCTG 3872
QY 611 CTATATCTGGCTTAAAGAAAGAGAGTCTTATGCGAGACTCACCTTTTCCAATCAACTCA 670
DB 3873 CTATATTTGGATAAAAAAAGCGAGGATTTATGCCGGAATTAACATTTTCTAATGAACATA 3932
QY 671 TCAGCAGAGATGAAGGAGCTTCACTGTGACTTTTCTGCTGCTGATGTTTCCCACTACTAGTAA 730
DB 3933 TAAGTAGACACGAAGGTTTACATTTGTGATTTTGGGTGTTTAAATGTTTAAACATTTATTAC 3992

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2005, 01:51:37 ; Search time 676.5 Seconds
(without alignments)
12855.457 Million cell updates/sec

Title: US-10-698-228-12

Perfect score: 1053

Sequence: 1 atggcgaccgcgaaggcc.....tcacctggtgcagatttt 1053

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9784742 seqs, 4129495052 residues

Total number of hits satisfying chosen parameters: 19569484

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
- 23: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 24: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 25: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
- 26: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 27: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 28: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1053	100.0	1053	19	US-10-698-228-12	Sequence 12, Appl
2	1053	100.0	1081	19	US-10-698-228-4	Sequence 4, Appl
3	1051.4	99.8	1053	19	US-10-698-228-2	Sequence 2, Appl
4	1051.4	99.8	4955	18	US-10-172-118-71	Sequence 71, Appl
5	1051.4	99.8	4955	19	US-10-342-887-71	Sequence 71, Appl

6	1051.4	99.8	4955	19	US-10-698-228-3	Sequence 3, Appl
7	580	55.1	2482	24	US-10-220-335-514	Sequence 524, App
8	580	55.1	2500	9	US-09-954-456-724	Sequence 724, App
9	580	55.1	2500	9	US-09-954-456-1169	Sequence 1169, Ap
10	580	55.1	2500	9	US-09-954-456-1827	Sequence 1827, Ap
11	580	55.1	2500	19	US-10-641-643-1370	Sequence 1370, Ap
12	580	55.1	2500	21	US-10-733-878-458	Sequence 458, App
13	580	55.1	2500	22	US-10-843-641A-3751	Sequence 3751, Ap
14	580	55.1	2500	22	US-10-843-641A-4196	Sequence 4196, Ap
15	580	55.1	2500	22	US-10-843-641A-4854	Sequence 4854, Ap
16	580	55.1	2500	24	US-10-756-149-713	Sequence 713, App
17	578.4	54.9	1989	9	US-09-925-301-505	Sequence 505, App
18	578.4	54.9	2216	16	US-10-084-817-342	Sequence 342, App
19	570.4	54.2	2113	22	US-10-764-420-1636	Sequence 1636, Ap
20	569	54.0	2641	24	US-10-220-335-170	Sequence 170, Appl
21	562.6	53.4	1328	19	US-10-403-571-75	Sequence 75, Appl
22	545.4	51.8	977	18	US-10-264-237-790	Sequence 790, App
23	514	48.8	1371	24	US-10-450-763-15278	Sequence 15278, A
24	463	44.0	1289	26	US-11-097-143-19181	Sequence 19181, A
25	447.8	42.5	186854	22	US-10-872-156-34	Sequence 34, Appl
26	442	42.0	1146	9	US-09-822-830A-174	Sequence 174, Appl
27	433.8	41.2	1218	16	US-10-128-714-7245	Sequence 7245, Ap
28	420.4	39.9	3945	26	US-11-097-143-19180	Sequence 19180, A
29	404.8	38.4	963	18	US-10-369-493-34052	Sequence 34052, A
30	398.6	37.9	1146	16	US-10-128-714-2245	Sequence 2245, Ap
31	394.2	37.4	1173	18	US-10-369-493-46252	Sequence 46252, A
32	393.6	37.4	1242	17	US-10-032-585-6537	Sequence 6537, Ap
33	388.2	36.9	1292	18	US-10-369-493-27809	Sequence 27809, A
34	382	36.3	1200	19	US-10-369-493-45857	Sequence 45857, A
35	380.6	36.1	1248	18	US-10-424-599-126843	Sequence 126843,
36	372.8	35.4	1361	19	US-10-424-599-40471	Sequence 40471, A
37	365.8	34.7	1586	19	US-10-424-599-73024	Sequence 73024, A
38	363.2	34.5	957	21	US-10-853-047-43	Sequence 43, Appl
39	363.2	34.5	1206	18	US-10-320-797-2145	Sequence 2145, Ap
40	350.6	33.3	1314	16	US-10-128-714-1245	Sequence 1245, Ap
41	350.6	33.3	1450	16	US-10-128-714-6245	Sequence 6245, Ap
42	350.6	33.3	3314	16	US-10-128-714-245	Sequence 245, App
43	350.6	33.3	3450	16	US-10-128-714-5245	Sequence 5245, Ap
44	349.8	33.2	629	24	US-10-450-763-6851	Sequence 6851, Ap
45	328.2	31.2	560	9	US-09-864-761-12381	Sequence 12381, A

ALIGNMENTS

RESULT 1

US-10-698-228-12
; Sequence 12, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WO0P
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-12

Query Match 100.0%; Score 1053; DB 19; Length 1053;

Best Local Similarity 100.0%; Pred. No. 9.3e-295;		Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	ATGGCGGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGATCATCTTCA	60
Db	1	ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGATCATCTTCA	60
Qy	61	GACACCAACGAAAGTGAATAAAGTCAATGAAGACCACTCTTAAGAAAGAGTCTCGC	120
Db	61	GACACCAACGAAAGTGAATAAAGTCAATGAAGACCACTCTTAAGAAAGAGTCTCGC	120
Qy	121	CGGTTTGTCACTTTCCAAATCCAGTACCTGATATTTGGAAATGATATAACAGGCACAG	180
Db	121	CGGTTTGTCACTTTCCAAATCCAGTACCTGATATTTGGAAATGATATAACAGGCACAG	180
Qy	181	GCTTCTCTTCGACACGACGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG	240
Db	181	GCTTCTCTTCGACACGACGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG	240
Qy	241	CTTAAAGCAGATGAGAAGTACTTCTCTCATCTTACATCTTAGCCCTTTTTCGAGCAGTGAT	300
Db	241	CTTAAAGCAGATGAGAAGTACTTCTCTCATCTTACATCTTAGCCCTTTTTCGAGCAGTGAT	300
Qy	301	GGAATTTGTAATGAAATTTGGTGGAGCGCTTTTGTAGCAGGTTCAGAGGCT	360
Db	301	GGAATTTGTAATGAAATTTGGTGGAGCGCTTTTGTAGCAGGTTCAGAGGCT	360
Qy	361	CGCTGTTCTATGGCTTTCAATTCATCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTG	420
Db	361	CGCTGTTCTATGGCTTTCAAAATTCATCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTG	420
Qy	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA	480
Db	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA	480
Qy	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGGATAGAGAGAA	540
Db	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGGATAGAGAGAA	540
Qy	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTGAAGAGGATTTCTTCTCAGGA	600
Db	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTGAAGAGGATTTCTTCTCAGGA	600
Qy	601	TCCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGCTTATGCCAGGACTCACCTTTTCC	660
Db	601	TCCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGCTTATGCCAGGACTCACCTTTTCC	660
Qy	661	AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTTCTTGCCTGATGTTCCAA	720
Db	661	AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTTCTTGCCTGATGTTCCAA	720
Qy	721	TACTTAGTAAATAGGCTTCAGAAAGAGGTTCAGGAGATCATTTGTGATGTCTCAAA	780
Db	721	TACTTAGTAAATAGGCTTCAGAAAGAGGTTCAGGAGATCATTTGTGATGTCTCAAA	780
Qy	781	ATTGAGCAGGATTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGATTCGATT	840
Db	781	ATTGAGCAGGATTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGATTCGATT	840
Qy	841	TTGATGAAACAGTACATTCAGTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTCTCA	900
Db	841	TTGATGAAACAGTACATTCAGTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTCTCA	900
Qy	901	AAGGTTTTTTCAGGAGAAATCCTTTTGAATTTTATGGAATAACATTTCTTTAGAGGAA	960
Db	901	AAGGTTTTTTCAGGAGAAATCCTTTTGAATTTTATGGAATAACATTTCTTTAGAGGAA	960
Qy	961	ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC	1020
Db	961	ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC	1020
Qy	1021	ACAGATAACGTCTTACCTTTGGATGAGATTTT 1053	

Db	1021	ACAGATAACGTCTTACCTTTGGATGAGATTTT 1053	
RESULT 2			
US-10-698-228-4			
; Sequence 4, Application US/10698228			
; Publication No. US20040072253A1			
; GENERAL INFORMATION:			
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA			
; TITLE OF INVENTION: New Protein and its DNA			
; FILE REFERENCE: 2619W00P			
; CURRENT APPLICATION NUMBER: US/10/698,228			
; CURRENT FILING DATE: 2003-10-30			
; PRIOR APPLICATION NUMBER: US/10/019,733			
; PRIOR FILING DATE: 2001-12-28			
; PRIOR APPLICATION NUMBER: JP 11-181131			
; PRIOR FILING DATE: 1999-06-28			
; PRIOR APPLICATION NUMBER: JP 11-192391			
; PRIOR FILING DATE: 1999-07-06			
; PRIOR APPLICATION NUMBER: JP 2000-017770			
; PRIOR FILING DATE: 2000-01-21			
; NUMBER OF SEQ ID NOS: 14			
; SEQ ID NO 4			
; LENGTH: 1081			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; OTHER INFORMATION:			
US-10-698-228-4			
Query Match 100.0%; Score 1053; DB 19; Length 1081;			
Best Local Similarity 100.0%; Pred. No. 9.4e-295;			
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	ATGGCGGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGATCATCTTCA	60
Db	20	ATGGCGGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGATCATCTTCA	79
Qy	61	GACACCAACGAAAGTGAATAAAGTCAATGAAGACCACTCTTAAGAAAGAGTCTCGC	120
Db	80	GACACCAACGAAAGTGAATAAAGTCAATGAAGACCACTCTTAAGAAAGAGTCTCGC	139
Qy	121	CGGTTTGTCTTCTTCCAAATCCAGTACCTGATATTTGGAAATGATATAACAGGCACAG	180
Db	140	CGGTTTGTCTTCTTCCAAATCCAGTACCTGATATTTGGAAATGATATAACAGGCACAG	199
Qy	181	GCTTCTCTTCGACACGACGAGGTTGACTTATCAAAAGGATCTCCCTCACTGGAACAAG	240
Db	200	GCTTCTCTTCGACACGACGAGGTTGACTTATCAAAAGGATCTCCCTCACTGGAACAAG	259
Qy	241	CTTAAAGCAGATGAGAAGTACTTCTCATCTCAGATCTTAGCCCTTTTTCGAGCAGTGAT	300
Db	260	CTTAAAGCAGATGAGAAGTACTTCTCATCTCAGATCTTAGCCCTTTTTCGAGCAGTGAT	319
Qy	301	GGAAATGTAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCAGAGGCT	360
Db	320	GGAAATGTAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCAGAGGCT	379
Qy	361	CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAAATGTTTATTAATGCAATTGAA	420
Db	380	CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAAATGTTTATTAATGCAATTGAA	439
Qy	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTAATGCAATTGAA	480
Db	440	CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTAATGCAATTGAA	499
Qy	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGGATAGAGAGAA	540
Db	500	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGGATAGAGAGAA	559
Qy	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGGATTTTCTTCTCAGGA	600
Db	560	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGGATTTTCTTCTCAGGA	619

QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAAAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
DB 620 TCTTTTGTCTATATTTCTGGCTAAAGAAAGAGAGGTCTTATGCCAGGACTCACTTTTTC 679
QY 661 AATGAACATCATCAGCAGAGATGAAGACTTCTACTGTGACTTTGGCTTGCCTGATGTTCCAA 720
DB 680 AATGAACATCATCAGCAGAGATGAAGACTTCTACTGTGACTTTGGCTTGCCTGATGTTCCAA 739
QY 721 TACTTAGTAAATAAGCTTTCAGAAAGAAAGGCTCAGGAGATCACTTGTGATGCTGTCAAA 780
DB 740 TACTTAGTAAATAAGCTTTCAGAAAGAAAGGCTCAGGAGATCACTTGTGATGCTGTCAAA 799
QY 781 ATTGACGAGGAGTTTAAACAGAAAGCTTGCAGGTTGGCTCATTTGGAATGAATTCGATT 840
DB 800 ATTGACGAGGAGTTTAAACAGAAAGCTTGCAGGTTGGCTCATTTGGAATGAATTCGATT 859
QY 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTTCTCA 900
DB 860 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTTCTCA 919
QY 901 AAGGTTTTTCAGGCGAGAAATCCTTTTGATTTTATGGAACAATTTCTTTAGAGGAAA 960
DB 920 AAGGTTTTTCAGGCGAGAAATCCTTTTGATTTTATGGAACAATTTCTTTAGAGGAAA 979
QY 961 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
DB 980 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1039
QY 1021 ACAGATAAGCTTTCACCTTGGATGCGAGATTTT 1053
DB 1040 ACAGATAAGCTTTCACCTTGGATGCGAGATTTT 1072

RESULT 3

US-10-698-228-2
; Sequence 2, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-2

Query Match 99.8%; Score 1051.4; DB 19; Length 1053;
Best Local Similarity 99.9%; Pred. No. 2.7e-294;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGGACCCGGAAGCGCGGAGCGGCTGGATCAGATCAGAGATCATCTTCA 60
DB 1 ATGGCGGACCCGGAAGCGCGGAGCGGCTGGATCAGATCAGAGATCATCTTCA 60
QY 61 GACACCAACGAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 61 GACACCAACGAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120

QY 121 CGGTTTCTCATCTTTTCCAAATCCAGTACCTCGATATTTGGAAAAATGTATAAACAGGACACAG 180
DB 121 CGGTTTCTCATCTTTTCCAAATCCAGTACCTCGATATTTGGAAAAATGTATAAACAGGACACAG 180
QY 181 GCTTCTTCTTGGCAGCAGAGAGGTTGACTTATCAAAAGGATCTCCCTCAGTGGAAACAAG 240
DB 181 GCTTCTTCTTGGCAGCAGAGAGGTTGACTTATCAAAAGGATCTCCCTCAGTGGAAACAAG 240
QY 241 CTTAAACGAGATGAGAAGTACTTCTCATCTCTCACATCTTGTAGCCTTTTGTGACGCCAGTGTAT 300
DB 241 CTTAAACGAGATGAGAAGTACTTCTCATCTCTCACATCTTGTAGCCTTTTGTGACGCCAGTGTAT 300
QY 301 GGAATTTGAAATGAAATTTGGTGGAGCGCTTTAGTTCAGAGAGTGCAGGTTCCAGAGGCT 360
DB 301 GGAATTTGAAATGAAATTTGGTGGAGCGCTTTAGTTCAGAGAGTGCAGGTTCCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCTGAGAATGTTCTCAGAGATGTCACAGTGTG 420
DB 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCTGAGAATGTTCTCAGAGATGTCACAGTGTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTTGAA 480
DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTTGAA 480
QY 481 ACCATGCCCTATGTTAAAGAAAAAGCAGATTTGGGCTTTCGATGGATAGCAGATAGAAAA 540
DB 481 ACCATGCCCTATGTTAAAGAAAAAGCAGATTTGGGCTTTCGATGGATAGCAGATAGAAAA 540
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTCTCTCTGTAGAAAGGAGTTTCTTCTCAGGA 600
DB 541 TCTACTTTTGGGAAAGAGTGGTGGCTTCTCTCTGTAGAAAGGAGTTTCTTCTCAGGA 600
QY 601 TCTTTTGTCTCTATATTTCTGGCTTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
DB 601 TCTTTTGTCTCTATATTTCTGGCTTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
QY 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
DB 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
QY 721 TACTTAGTAAATAAGCTTTCAGAAAGAAAGGCTCAGGAGAGATCACTTGTGATGCTGTCAAA 780
DB 721 TACTTAGTAAATAAGCTTTCAGAAAGAAAGGCTCAGGAGAGATCACTTGTGATGCTGTCAAA 780
QY 781 ATTGACGAGGAGTTTAAACAGAAAGCTTGCAGTTCGCTCATTTGGAATGAATTCGATT 840
DB 781 ATTGACGAGGAGTTTAAACAGAAAGCTTGCAGTTCGCTCATTTGGAATGAATTCGATT 840
QY 841 TTGATGAACAGTACATTTGAGTTTGTAGCTCAGCAGATTACTTGTGGAACCTTGGATTTCTCA 900
DB 841 TTGATGAACAGTACATTTGAGTTTGTAGCTCAGCAGATTACTTGTGGAACCTTGGATTTCTCA 900
QY 901 AAGGTTTTTTCAGGCGAGAAATCCTTTTGAATTTTATGGAACAATTTCTTTAGAAAGGAAA 960
DB 901 AAGGTTTTTTCAGGCGAGAAATCCTTTTGAATTTTATGGAACAATTTCTTTAGAAAGGAAA 960
QY 961 ACAAAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTGTATGCGAGAAACC 1020
DB 961 ACAAAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTGTATGCGAGAAACC 1020
QY 1021 ACAGATAAGCTTTCACCTTGGATGCGAGATTTT 1053
DB 1021 ACAGATAAGCTTTCACCTTGGATGCGAGATTTT 1053

RESULT 4

US-10-172-118-71
; Sequence 71, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter

APPLICANT: Mao, Mao
APPLICANT: Roberts, Chris
APPLICANT: Van 't Veer, Laura
APPLICANT: Van de Vijver, Marc
APPLICANT: Bernards, Rene
TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
FILE REFERENCE: 9301-175-999
CURRENT FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: 60/380,770
PRIOR FILING DATE: 2002-05-14
NUMBER OF SEQ ID NOS: 2699
SEQ ID NO 71
LENGTH: 4955
TYPE: DNA
ORGANISM: Homo sapiens
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: AB036063
DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-71

Query Match 99.8%; Score 1051.4; DB 18; Length 4955;
Best Local Similarity 99.9%; Pred. No. 6.8e-294;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 245 ATGGGCGACCCGGAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCCTGC 120
DB 305 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCCTGC 364
QY 121 CGTTTGTCTATCTTTCGAATCAGTACCTTGATTTGGAAATGATTAACAGGCACAG 180
DB 365 CGTTTGTCTATCTTTCGAATCAGTACCTTGATTTGGAAATGATTAACAGGCACAG 424
QY 181 GCTTCTCTTGGACAGCAGAGAGTGTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
DB 425 GCTTCTCTTGGACAGCAGAGAGTGTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 484
QY 241 CTAAAGCAGATGAGAGTACTTCTATCTCTACATCTTACGCTTTTTCAGAGCAGTGAT 300
DB 485 CTAAAGCAGATGAGAGTACTTCTATCTCTACATCTTACGCTTTTTCAGAGCAGTGAT 544
QY 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAAGAGTGCAGGTTCCAGAGGCT 360
DB 545 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAAGAGTGCAGGTTCCAGAGGCT 604
QY 361 CGCTGTTCTTATGGCTTTCAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 420
DB 605 CGCTGTTCTTATGGCTTTCAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAATTTTATTTAATGCAATGAA 480
DB 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAATTTTATTTAATGCAATGAA 724
QY 481 ACATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGCATGATGAGATAGAGAGAA 540
DB 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGCATGATGAGATAGAGAGAA 784
QY 541 TCTACTTTTGGGAAAGAGTGTGGGCTTGTGCTGTAGAGAGAGTTCCTCTCAGGA 600
DB 785 TCTACTTTTGGGAAAGAGTGTGGGCTTGTGCTGTAGAGAGAGTTCCTCTCAGGA 844
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGAGTCTTATGCGAGAGTCTCCTTTTTC 660
DB 845 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGAGTCTTATGCGAGAGTCTCCTTTTTC 904
QY 661 AATGAACTCATCAGCAGAGATGAAGACCTTCACTGTGACTTTTGTGCTGATGTTCCAA 720
DB 905 AATGAACTCATCAGCAGAGATGAAGACCTTCACTGTGACTTTTGTGCTGATGTTCCAA 964

QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTCAGGAGAGATCATTTGTGATGCTGTCAA 780
DB 965 TACTTAGTAAATAAGCCTTCAGAAAGAGGTCAGGAGAGATCATTTGTGATGCTGTCAA 1024
QY 781 ATTGAGCAGAGTATTTTAAACAGAGCCTTGCAGTTCGCTCATTTGGAATGAATGCAAT 840
DB 1025 ATTGAGCAGAGTATTTTAAACAGAGCCTTGCAGTTCGCTCATTTGGAATGAATGCAAT 1084
QY 841 TTGATGAACACAGTACATTTGAGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATCTCA 900
DB 1085 TTGATGAACACAGTACATTTGAGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATCTCA 1144
QY 901 AAGTATTTTTCAGGCAGAAAAATCCTTTTGAATTTTATGAAAAACATTTCTTTAGAGAAAA 960
DB 1145 AAGTATTTTTCAGGCAGAAAAATCCTTTTGAATTTTATGAAAAACATTTCTTTAGAGAAAA 1204
QY 961 ACAATTTTCTTTCAGAAACAGATTTTCAGATATCAGCGTTTTCAGATTTATGCGAGAAACC 1020
DB 1205 ACAATTTTCTTTCAGAAACAGATTTTCAGATATCAGCGTTTTCAGATTTATGCGAGAAACC 1264
QY 1021 ACAGATAACGTCCTTCACCTTGGATGCAAGATTTT 1053
DB 1265 ACAGATAACGTCCTTCACCTTGGATGCAAGATTTT 1297

RESULT 5
US-10-342-887-71
; Sequence 71, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-71

Query Match 99.8%; Score 1051.4; DB 19; Length 4955;
Best Local Similarity 99.9%; Pred. No. 6.8e-294;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 245 ATGGGCGACCCGGAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCCTGC 120
DB 305 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCCTGC 364
QY 121 CGTTTGTCTATCTTTCGAATCAGTACCTTGATTTGGAAATGATTAACAGGCACAG 180
DB 365 CGTTTGTCTATCTTTCGAATCAGTACCTTGATTTGGAAATGATTAACAGGCACAG 424
QY 181 GCTTCTCTTGGACAGCAGAGAGTGTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240

Db 425 GCTTCCTCTCGACAGAGAGGTCGACTTATCAAGAGATCTCCCTCACTGGAACAAG 484
Qy 241 CTTAAAGCAGATGAGAGTACTTCACTCTCAATCTTAGCCTTTTTCAGCAGTAT 300
Db 485 CTTAAAGCAGATGAGAGTACTTCACTCTCAATCTTAGCCTTTTTCAGCAGTAT 544
Qy 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTTCAGAGGCT 360
Db 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTTCAGAGGCT 604
Qy 361 CGCTGTTCTATGGCTTTCAATCTCATCGAGAAATGTTTCACTCAGAGATGTCAGTTG 420
Db 605 CGCTGTTCTATGGCTTTCAATCTCATCGAGAAATGTTTCACTCAGAGATGTCAGTTG 664
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATGAA 724
Qy 481 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGCCCTTGGATGGATAGCAGATAGAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGCCCTTGGATGGATAGCAGATAGAAA 784
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAAGAGAGTTCCTTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAAGAGAGTTCCTTCTCAGGA 844
Qy 601 TCTTTTGTCTATATCTGGCTTAAGAAAGAGAGTTCCTTATGCGAGACTCCTTTTCC 660
Db 845 TCTTTTGTCTATATCTGGCTTAAGAAAGAGAGTTCCTTATGCGAGACTCCTTTTCC 904
Qy 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTTCTTGGCTGATGTTCCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTTCTTGGCTGATGTTCCAA 964
Qy 721 TACTTAGTAAATAAGCTTTAGAGAAAGGTCAGGAGATCATTTGTAAGTGTCTCAAA 780
Db 965 TACTTAGTAAATAAGCTTTAGAGAAAGGTCAGGAGATCATTTGTAAGTGTCTCAAA 1024
Qy 781 ATTGACGAGAGTTCCTTCAAGAGCTTTGCCAGTTCCTTCAATGGAATGCAATT 840
Db 1025 ATTGACGAGAGTTCCTTCAAGAGCTTTGCCAGTTCCTTCAATGGAATGCAATT 1084

RESULT 6

US-10-698-228-3
; Sequence 3, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28

; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 3
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-3

Query Match 99.8%; Score 1051.4; DB 19; Length 4955;

Best Local Similarity 99.9%; Pred. No. 6.8e-294;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGCGACCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 245 ATGGCGACCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
Qy 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCACCTCTTAAGAAAGAGTTCGCG 120
Db 305 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCACCTCTTAAGAAAGAGTTCGCG 364
Qy 121 CGCTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAATGTATAACAGGCACAG 180
Db 365 CGCTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAATGTATAACAGGCACAG 424
Qy 181 GCTTCCTCTTGGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db 425 GCTTCCTCTTGGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 484
Qy 241 CTTAAAGCAGATCAGAGTACTTCACTCTCAATCTTACGCTTTTTCAGGCCAGTCAAT 300
Db 485 CTTAAAGCAGATCAGAGTACTTCACTCTCAATCTTACGCTTTTTCAGGCCAGTCAAT 544
Qy 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTTCAGAGGCT 360
Db 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTTCAGAGGCT 604
Qy 361 CGCTGTTTCTATGCTTTCAAAATTTCTCATCGAGATGTTCACTCAGAGATGTCAGAGTTG 420
Db 605 CGCTGTTTCTATGCTTTCAAAATTTCTCATCGAGATGTTCACTCAGAGATGTCAGAGTTG 664
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATGAA 724
Qy 481 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGCCCTTGGATGGATAGCAGATAGAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGCCCTTGGATGGATAGCAGATAGAAA 784
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAAGAGAGTTCCTTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAAGAGAGTTCCTTCTCAGGA 844
Qy 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGGTCCTTATGCCAGGACTCCTTTTTC 660
Db 845 TCTTTTGTCTATATTTCTGGCTAAAGAGAGGTCCTTATGCCAGGACTCCTTTTTC 904
Qy 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTTCTTGGCTGATGTTCCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTTCTTGGCTGATGTTCCAA 964
Qy 721 TACTTAGTAAATAAGCTTTAGAGAAAGGTCAGGAGATCATTTGTAAGTGTCTCAAA 780
Db 965 TACTTAGTAAATAAGCTTTAGAGAAAGGTCAGGAGATCATTTGTAAGTGTCTCAAA 1024
Qy 781 ATTGACGAGAGTTCCTTCAAGAGCTTTGCCAGTTCCTTCAATGGAATGCAATT 840
Db 1025 ATTGACGAGAGTTCCTTCAAGAGCTTTGCCAGTTCCTTCAATGGAATGCAATT 1084

; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 724
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-724

Query Match
Best Local Similarity 55.1%; Score 580; DB 9; Length 2500;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 89 ATCAAGAGCCACTCTAAGAAAGATTCCTCGCGGTTGTGTCATCTTCCATCCAGTACC 148
DB 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTGTCATCTTCCCATCCAGTACC 456
QY 149 CTGATATTTGGAAATGTATAACAGCAGCAGGCTTCTTCTGACAGCAGAGAGGTTG 208
DB 457 ATGATATCTGCGAGATGTATAAGAGCAGAGGCTTCTTCTTGACCGCGAGAGGTTG 516
QY 209 ACTTATCAAGAGTCTCCCTCACTGGAACAAAGCTTAAAGCAGATGAGAAGTACTTCATCT 268
DB 517 ACCTCTCAAGAGCATTTCAGCAGCTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
QY 269 CTCACATCTTAGCCCTTTTTCGAGCCAGTGAATGTAATGTAATGTAATGTAATGTAATG 328
DB 577 CCCATGTTCTGGCTTTCTTTGACAGCAAGCGATGATGATGATGATGATGATGATGATG 636
QY 329 GCTTTAGTCAGAGGTCAGGTTCCAGAGGCTCGCTGTTTCTATGCTTCCAAATTCCTCA 388
DB 637 GATTTAGCAAGAGTTTCAGATTCAGAAAGCCGCTGTTTCTATGCTTCCAAATTCCTCA 696
QY 389 TCAGAGATGTTCACTCAGAGATGTACAGTTTGTGCTGATAGACACTTACATCAGAGATCCCA 448
DB 697 TGGAAACATACATCTGAAAGTATGATGCTTCTTATGACATTTACATAAAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTAATGCAATGAAACCAATGCAATGCAATGCAATGCAATGCAATG 508
DB 757 AAGAAAGGGAATTTCTCTTCAATGCAATGAAACCAATGCAATGCAATGCAATGCAATGCAATG 816
QY 509 ATTGGGCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 568
DB 817 ACTGGGCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 876
QY 569 TTGCTGCTAGAGAGGATTTTCTTCTCAGGATCTTTTGTGCTGATATTTCTGCTAAAGA 628
DB 877 TTGCTGCTAGAGAGGATTTTCTTCTCAGGATCTTTTGTGCTGATATTTCTGCTCAAGA 936
QY 629 AGAGAGGCTTATGCGAGGACTCATTCTTCCATGAACCTCATCAGCAGAGATGAAGGAC 688
DB 937 AACGAGGACTGATGCTGCGCTCACATTTCTTAATGACTTATAGCAGATGAGGTT 996
QY 689 TTCACGTGACATTTGCTGCTGATGTTTCCAAATCTTATGATGATGATGATGATGATGATG 748
DB 997 TACACTGTGATTTGCTGCTGATGTTTCAAAACACCTGCTGATGATGATGATGATGATGATG 1056
QY 749 GGTCAAGGAGATCATTTGATGCTGCTCAAAATTCAGCAGGAGTTTAAACAGAGGCT 808
DB 1057 GAGTAAGAAATAATATCAATGCTGCTGCGGATGAAACAGGAGTTTCTCAGTGGGCT 1116
QY 809 TGCAGTTGGCTCATTTGGAATGAAATGCAATTTTGAAGAAACAGTACATTTGATGTTG 868
DB 1117 TGCTGTGAGCTCATTTGGAATGAAATGCAATTTTGAAGAAACAGTACATTTGATGTTG 1176
QY 869 CTGACAGATTTACTGTGGAATTTGGAATTTCTCAAGGTTTTCAGGAGGAAATCTTTTG 928
DB 1177 CAGACAGACTTATGCTGGAATTTGGAATTTCTCAAGGTTTTCAGGAGGAAATCTTTTG 1236
QY 929 ATTTATGGAACATTTCTTATAGAGGAAACAAATTTCTTTGAGAAACAGATTTTCAG 988

DB 1237 ACTTTATGAGAAATATTTTCACTGGAAGAAAGACTTAATCTTCTTGAGAAGAGAGTAGCG 1296
QY 989 AGTATCAGCGTTTTCAGTTATGCGAGAAACCCAGAGATAACGCTCTTCCACCTTGGATGAG 1048
DB 1297 AGTATCAGAGATGGGAGTGTATGTCAGAGTCCACAGAGAAATCTTTTACCTTTGGATGCTG 1356
QY 1049 ATTT 1052
DB 1357 ACTT 1360

RESULT 9
US-09-954-456-1169
; Sequence 1169, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1169
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1169

Query Match
Best Local Similarity 55.1%; Score 580; DB 9; Length 2500;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 89 ATCAAGAGCCACTCTAAGAAAGATTCCTCGCGGTTGTGTCATCTTCCATCCAGTACC 148
DB 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTGTCATCTTCCCATCCAGTACC 456
QY 149 CTGATATTTGGAAATGTATAACAGCAGCAGGCTTCTTCTGACAGCAGAGAGGTTG 208
DB 457 ATGATATCTGCGAGATGTATAAGAGCAGAGGCTTCTTCTTGACCGCGAGAGGTTG 516
QY 209 ACTTATCAAGAGTCTCCCTCACTGGAACAAAGCTTAAAGCAGATGAGAAGTACTTCATCT 268
DB 517 ACCTCTCAAGAGCATTTCAGCAGCTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
QY 269 CTCACATCTTAGCCCTTTTTCGAGCCAGTGAATGTAATGTAATGTAATGTAATGTAATG 328
DB 577 CCCATGTTCTGGCTTTCTTTGACAGCAAGCGATGATGATGATGATGATGATGATGATG 636
QY 329 GCTTTAGTCAGAGGTCAGGTTCCAGAGGCTCGCTGTTTCTATGCTTCCAAATTCCTCA 388
DB 637 GATTTAGCAAGAGTTTCAGATTCAGAAAGCCGCTGTTTCTATGCTTCCAAATTCCTCA 696

Db 1177 CAGACAGACTTATGCTGGAACCTGGTGTAGCAAGGTTTTCAGAGTAGAGAACCCATTTG 1236
QY 929 ATTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAACGATTTTCAG 988
Db 1237 ACTTTATGAGAAATATTTCACTGGAAGGAAAGACTAACTTTCTTTGAGAGAGAGTAGGCG 1296
QY 989 AGTATCAGCGTTTTCAGTTATGGCAGAAACACACAGATAACGCTTTCACCTTCGATGCGAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGATGTCAGTCCACAGAGAAATTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 11

US-10-641-643-1370
; Sequence 1370, Application US/10641643
; Publication No. US20040077003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; Susan G. Stuart
; Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
; GENE EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/641,643
FILING DATE: 14-Aug-2003

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: <Unknown>

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0001 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 1370:

SEQUENCE CHARACTERISTICS:

LENGTH: 2500 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GENBANK

CLONE: g36154

SEQUENCE DESCRIPTION: SEQ ID NO: 1370 :

US-10-641-643-1370

Query Match 55.1%; Score 580; DB 19; Length 2500;

Best Local Similarity 75.1%; Pred. No. 4.6e-157; Mismatches 240; Indels 0; Gaps 0;

Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGCGGGTTTGTCTATCTTTTCAATCCAGTACC 148

Db 397 AGGATGAGCGGCTGCTGAGAGAAACCCCGCGGCTTGTCTATCTTTCCCATCGAGTACC 456

QY 149 CTGATATTTGGAAAAATGTAATAACAGGCACAGGCTTCTCTTGGACAGCAGAGAGGTTG 208
Db 457 ATGATATCTGGCAGATGTAATAAGAGGACAGAGGCTTCTCTTTTGGACCCCGAGAGGTTG 516
QY 209 ACTTATCAAAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGAGATGAGAAGTACTTCACT 268
Db 517 ACCTCTCAAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
QY 269 CTCACTATTTAGGCTTTTTCAGCCAGTGAATGTAATGTAATGTAATGTAATGTAATGTAATG 328
Db 577 CCCATGTTCTGGCTTTCTTTGACAGCAAGCGATGATGTAATGTAATGTAATGTAATGTAATG 636
QY 329 GCTTTAGTCAGGAGTGCAGGTTCCAGAGGCTCCCTGTTCTATGGCTTTTCAAAATTTCTCA 388
Db 637 GATTTAGCAAGAGTTTCAGATTACAGAGCCGCTGTTCTTCTATGGCTTCCAAATTTGCCA 696
QY 389 TCGAGAAATGTTCACTCAGAGATGTPACAGATTTGCTGTATGATAGACACTTTACATCAGAGATCCCA 448
Db 697 TGGAAAAACATACATTTCTGAAATGTAATGTTCTTTTATGACACTTTACATAAAAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCCCTATGTTAAGAAAAAGCAG 508
Db 757 AAGAAAGGGAATTTCTCTCAATGCCATTTGAAACGATGCCCTGTTGTTCAAGAAAGAGGCG 816
QY 509 ATTGGGCTTTGCGATGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCT 568
Db 817 ACTGGGCTTTGCGCTGATTTGGGACAAAGAGGCTACTATGGTGAACGTTGTTAGCCT 876
QY 569 TTGCTGCTGTAGAAGGAGTTTCTTCTCAGAGATCTTTTGTGCTATATTTGGCTTAAAGA 628
Db 877 TTGCTGCAAGTGAAGGCAATTTCTTTTCGCTGCTTTTTCGCTGCTATTTTGGCTCAAGA 936
QY 629 AGAGAGTCTTATGCCAGGACTCACTTTTTCATGAACTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCCCTGGCTCACATTTTCTAATGAACCTTTATAGCAGAGATGAGGTT 996
QY 689 TTCACTGTGACTTTGCTTGGCTGATGTTCCAACTTCTAGTAAATAGCCTTTCAGAGGAA 748
Db 997 TACACTGTGATTTTGTCTGCTGATGTTCAAAACACCTGGGTACACAAACCATCGGAGGAGA 1056
QY 749 GGGTCAGGAGATCATTTGTTGATGCTCTCAAAATTTGAGCAGGAGTTTTAAACAGAGGCT 808
Db 1057 GAGTAAAGAGAAATTAATCAATGCTGTTCCGATAGAACAGGAGTTCTCTACTAGGCT 1116
QY 809 TGCAGTTGGCCTCATTTGGAATGAATTTGCAATTTTGTGAAACAGTACATGAGTTTGTAG 868
Db 1117 TGCCTGTGAAGCTCATTTGGGATGAAATGCACTCTAATGAAGCAATACATTCAGTTTGTGG 1176
QY 869 CTGACAGATTACTTGTGGAACCTTGGATCTCAAGGTTTTCAGGCGAGAAATCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACCTGGGTTTTCAGCAAGGTTTTCAGAGTAGAGAACCCATTTG 1236
QY 929 ATTTTATGAAAAACATTTCTTTAAGAGGAAAAACAAATTTCTTTGAGAAAAACGAGTTTCAG 988
Db 1237 ACTTTATGAGAAATATTTCTGGAAGGAAAGACTAACTTTCTTTGAGAGAGAGTAGGCG 1296
QY 989 AGTATCAGCGTTTTCAGAGTTATGGCAGAAACCCAGATAACGCTTCTTCACTTGGATGAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGTGTCAGTCCACAGAGAAATTTCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 12

US-10-733-878-458
; Sequence 458, Application US/10733878
; Publication No. US20040224408A1
; GENERAL INFORMATION:
; APPLICANT: Jean-Philippe Girard
; APPLICANT: Francois Amalric
; APPLICANT: Myriam Roussigne

APPLICANT: Thomas Clouaire
TITLE OF INVENTION: THAP PROTEINS AS NUCLEAR RECEPTORS FOR
TITLE OF INVENTION: CHEMOKINES AND ROLES IN TRANSCRIPTIONAL REGULATION, CELL
TITLE OF INVENTION: PROLIFERATION AND CELL DIFFERENTIATION
FILE REFERENCE: BIOBANK.012A
CURRENT APPLICATION NUMBER: US/10/733,878
CURRENT FILING DATE: 2003-12-10
PRIOR APPLICATION NUMBER: 60/432699
PRIOR FILING DATE: 2002-12-10
PRIOR APPLICATION NUMBER: 60/485027
PRIOR FILING DATE: 2003-07-03
NUMBER OF SEQ ID NOS: 535
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 458
LENGTH: 2500
TYPE: DNA
ORGANISM: Homo sapiens
US-10-733-878-458

Query Match 55.1%; Score 580; DB 21; Length 2500;
Best Local Similarity 75.1%; Pred. No. 4.6e-157;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCACTCTTAAGAAAGAGTTCTGCGCGTTTGTCTCATCTTTCCAAATCCAGTACC 148
DB 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTCTCATCTTTCCCGCATCGAGTACC 456
QY 149 CTGATATTTGGAAATGTATAACAGGACAGGCTTCTTGGACAGAGAGGTTG 208
DB 457 ATGATATCTGGCAGATGTATAAGAGGACAGAGGCTTCTTTTGGACCGCGAGGTTG 516
QY 209 ACTTATCAAGAGATCTCCCTCACTGAGCAAGCTTAAAGCAGATGAGAGTACTTCATCT 268
DB 517 ACTCTCCAGACATTCAGACTGGGAATCCCTGAACCCGAGGAGAGATATTTATAT 576
QY 269 CTCACATCTTAGCCTTTTTCAGCGAGTGATGGAATTTGAAATGAAATTTGGTGGAGC 328
DB 577 CCCATGTTCTGCTTCTTTTGCAGCAAGCATGTCATGTAATGAATGAATTTGGTGGAGC 636
QY 329 GCTTTAGTCAGAGGTGAGGTTCCAGAGGCTGCTGTTTCTATGGCTTTCAAATTTCTCA 388
DB 637 GATTTAGCCAAAGAGTTTCAGATTACAGAGCCGCTGTTTCTATGGCTTTCCAAATTTGCCA 696
QY 389 TCGAGAACTTCACTCAGAGAGTACAGTTTGTGTATAGACACTTACATCAGAGATCCCA 448
DB 697 TGAAGAAACATACATTCGAAATGTATAGTCTTTTATGACACTTACATAAAGATCCCA 756
QY 449 AGAAAGGGAAATTTTATTTAATGCAATTTGAACCATGCCCTATGTTAAGAAAAAGCAG 508
DB 757 AAGAAAGGAAATTTCTTCAATGCCATTTGAACGATGCTTGTGTCAAGAGAGGCGAG 816
QY 509 ATTGGGCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 568
DB 817 ACTGGGCTTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 876
QY 569 TTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 628
DB 877 TTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 936
QY 629 AGAGAGGCTTATGCGAGGACTCACATTTTTCOAATGAACTCATCAGCAGATGAAGGAC 688
DB 937 AACGAGGACTGATGCTGCGCTCACATTTTCTAATGAACTTATGACAGATGAGGTT 996
QY 689 TTTACTGTGACTTTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 748
DB 997 TACACTGTGATTTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1056
QY 749 GGGTTCAGGAGATCATTTGATGCTGTCAAAATTCAGCAGAGGATTTTAAACAGAGCCT 808
DB 1057 GAGTAAGAGAAATAATTAATTAATGCTGTTTCGATGATGATGATGATGATGATGATGAT 1116
QY 809 TSCAGTTGCGCTCATTTGGAATGAATTTGATGATGATGATGATGATGATGATGATGAT 868

DB 1117 TGCCTGTGAAGCTCATTTGGATGAATTCGACTCTAATGAAGCAATACATTGAGTTGTGG 1176
QY 869 CTGACAGATTACTTGTGGAACCTTGGATTTCTCAAGAGGTTTTCAGGCGAGAAATCCTTTTG 928
DB 1177 CAGACAGACTTATGCTGGAACCTGCTTTAGCAAGGTTTTCAGAGTAGAGACCAATTTG 1236
QY 929 ATTTATGGAACCATTTCTTTTGAAGGAAAAACAAATTTCTTTTGAGAAACGAGTTTCAG 988
DB 1237 ACTTTATGAGAATATTTCACTGGAAGGAAGAACTAACTTTTTCAGAGAGAGTAGGCG 1296
QY 989 AGTATCAGCGTTTTCAGTATGAGCAGAAACCAAGATACGTTCTTCACTTGGATGAGCAG 1048
DB 1297 AGTATCAGAGAGTGGGAGTATGCTCAAGTCCAAACAGAGAAATTTCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
DB 1357 ACTT 1360

RESULT 13

US-10-843-641A-3751
Sequence 3751, Application US/10843641A
Publication No. US2005006454A1
GENERAL INFORMATION:
APPLICANT: Avalon Pharmaceuticals, Inc.
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
TITLE OF INVENTION: Signature Gene Sets
FILE REFERENCE: 689290-189
CURRENT APPLICATION NUMBER: US/10/843,641A
CURRENT FILING DATE: 2004-05-12
PRIOR APPLICATION NUMBER: US/09/873,367
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US/09/954,531
PRIOR FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/09/954,456
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US/09/962,436
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US/09/962,832
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US/09/964,824
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: US/09/967,768
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US/09/968,007
PRIOR FILING DATE: 2001-10-02
PRIOR APPLICATION NUMBER: US/09/969,347
PRIOR FILING DATE: 2001-10-02
PRIOR APPLICATION NUMBER: US/09/969,708
PRIOR FILING DATE: 2001-10-03
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 8447
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3751
LENGTH: 2500
TYPE: DNA
ORGANISM: Homo sapiens
US-10-843-641A-3751

Query Match 55.1%; Score 580; DB 22; Length 2500;
Best Local Similarity 75.1%; Pred. No. 4.6e-157;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCACTCTTCAAGAAAGAGTTCTGCGCGTTTGTCTCATCTTTCCAAATCCAGTACC 148
DB 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTCTCATCTTTCCCGCATCGAGTACC 456
QY 149 CTGATATTTGGAAATGTATAACAGGACAGGCTTCTTCTGACAGCAGAGAGGTTG 208
DB 457 ATGATATCTGGCAGATGTATAAGAGGACAGAGGCTTCTTTTGGACCGCGAGGAGGTTG 516
QY 209 ACTTATCAAGAGATCTCCCTCACTGGAACAAAGCTTAAAGCAGATGAGAGTACTTCATCT 268

Db 517 ACCTCTCCAGGACATTCAGCATCTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
Qy 269 CTCACATCTTAGCCTTTTTCAGCCAGCATGGAATGTAAATGAAATTTGGTGAGC 328
Db 577 CCCATGTTCTGGCTTTCTTTGACGACGATGCGCATAGTAATGAAACTTTGGTGAGC 636
Qy 329 GCTTTAGTCAGAGGTCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTTCA 388
Db 637 GATTTAGCCAAGAAAGTTTCAGATTACAGAACCCGCTGTTTCTATGGCTTCAAATTTG 696
Qy 389 TCGAGATCTTCACACAGAGATGACAGTTTGTGTAGACATACATCAGAGATCCCA 448
Db 697 TGAAGAACATACATCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 756
Qy 449 AGAAAGGGAATTTTATTTAATGCAATGMAACCATGCTCTATGTTAAGAAAAGCAG 508
Db 757 AGAAGGGAATTTCTCTCAATGCAATGMAACCATGCTCTATGTTAAGAAAAGCAG 816
Qy 509 ATTGGGCTTGGCATGAGATGACAGATGAGAAATCTACTTTTGGGAAAGAGTGCTGCT 568
Db 817 ACTGGGCTTGGCTGATTTGGGACAAAGAGGCTACCTATGCTGAACGTTGTAGCCT 876
Qy 569 TTGCTGCTAGAGAGGTTTCTCTCAGGATCTTTTGTGTCTATATTTCTGGCTAAGA 628
Db 877 TTGCTGCTAGAGAGGCTTTCTTTTCCGGTCTTTTGTGCTGATATTTCTGGCTCAAG 936
Qy 629 AGAGAGGCTTATGCGAGGACTCACATTTTCCAAATGCAATCTCAGCAGAGATGAGGAC 688
Db 937 AACGAGGACTGATGCTGGGCTTCAATTTTCTAATGAACTTATTTAGCAGAGATGAGGTT 996
Qy 689 TTCACTGTGACTTTTGTGCTGATGTTCCAAATGCTTAAATAGTAAATAGCCTTCAAGAAA 748
Db 997 TACACTGTGATTTTGTGCTGATGTTTCAACACCTGGTACACAAACCATTCGGAGGAGA 1056

RESULT 14

US-10-843-641A-4196
; Sequence 4196, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531

; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4196
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-4196

Query Match 55.1%; Score 580; DB 22; Length 2500;
Best Local Similarity 75.1%; Pred. No. 4.6e-157;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
Qy 89 ATGAAGAGCCACTCTCTAAGAAAGATTCTCGCCGTTTGTCTATCTTTTCCAATCCAGTACC 148
Db 397 AGATGAGCCGCTGCTGAGAGAAACCCGCCCTTTGTCTATCTTTCCCATCCAGTACC 456
Qy 149 CTGATATTTGAAATATATAAAGCAGCAGCAGGCTTCCCTCTGGACAGCAGAGAGGTTG 208
Db 457 ATGATATCTGCAGATGTATAAGAGGAGAGGCTTCTTTTGGACCCCGAGAGGTTG 516
Qy 209 ACTTATCAAGAGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAGTACTTCTATCT 268
Db 517 ACCTCTCAAGAGCAATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
Qy 269 CTCACATCTTAGCCTTTTTCAGCCAGTGTGAATGGAATGTAATGAAATTTGGTGAGC 328
Db 577 CCATGTTCTGGCTTTCTTTGACAGCAAGCGATGCAATGAAATGAAATTTGGTGAGC 636
Qy 329 GCTTTAGTCAGGAGGTCAGGTTCCAGAGGCTCCCTGTTCTATGGCTTTTCAAAATTTCTCA 388
Db 637 GATTTAGCCAAGAGTTTCAGATTACAGAGCCGCTGTTCTATGGCTTCCAAATTTGCCA 696
Qy 389 TCAGAGATGTTCACTCAGAGATGTACAGTTTGTGTGATAGACACTTACATCAGAGATCCCA 448
Db 697 TGAAGAACATACATCTCTGAAATGTATAGTCTTCTTATTTGACACTTACATAAAGATCCCA 756
Qy 449 AGAAAGGGAATTTTATTTAATGCAATTCGAACCATGCTCTATGTTAAGAAAAGCAG 508
Db 757 AGAAAGGGAATTTTCTCTCAATGCAATTCGAACCATGCTCTGTTGTCAGAGAGAGGAGC 816
Qy 509 ATTGGGCTTGGCATGAGATGAGAAATCTACTTTTGGGAAAGAGTGTTGGCT 568
Db 817 ACTGGGCTTGGCTGATTTGGGACAAAGAGGCTACTATGGTGAACGTTGTAGCCT 876
Qy 569 TTGCTGCTGAGAGGAGTTTCTTCTCAGGATCTTTTGTGCTGATATTTCTGGCTAAGA 628
Db 877 TTGCTGAGTGAAGGAGTTTCTTTTCCGGTCTTTTGTGCTGATATTTCTGGCTCAAGA 936
Qy 629 AGAGAGGCTTATGCGAGGACTCACATTTTCCAAATGCAATCTCAGCAGAGATGAGGAC 688
Db 937 AACGAGGACTGATGCTGGGCTTCAATTTTCTAATGAACTTATTTAGCAGAGATGAGGTT 996
Qy 689 TTCACTGTGACTTTTGTGCTGATGTTCCAAATGCTTAAATAGTAAATAGCCTTCAAGAAA 748
Db 997 TACACTGTGATTTTGTGCTGATGTTTCAACACCTGGTACACAAACCATTCGGAGGAGA 1056

QY 749 GGGTCAGGAGATCATTTGTTGCTGCTCAAAATTCAGCAGGAGTTTAAACAGAGCCT 808
DB 1057 GAGTAAGAGAAATTAATCAATGCTGTTCCGATAGAACAGGAGTTCCCTACAGAGCCT 1116
QY 809 TGCCAGTTGGCCCTCAITGGAAATGCAITTTTGATGAAACAGTACATTTGAGTTTGTAG 868
DB 1117 TGCCGTGTAAGCTCATTTGGGATGAATTCACCTCAATGAAGCAATACATTTGAGTTGTG 1176
QY 869 CTGACAGATTAATTTGTTGGAACCTTGAATTCCTCAAGGTTTTCAGGAGAAAATCCCTTTG 928
DB 1177 CAGACAGACTTATGCTGGAACCTGGGTTTATAGCAAGGTTTTCAGAGTAGAGAACCCATTG 1236
QY 929 ATTTTATGAAACATATCTTTAGAGGAAACCAATTTCTTTGAGAAACGAGTTTCAG 988
DB 1237 ACTTTATGAGAAATATTTCACTGGAAGAAAGACTAACTTTCTTTGAGAGAGTAGGCG 1296
QY 989 AGTATCAGGCTTTTGCACTTATGGCAGAAACACAGATAACGCTTTCACCTTGGATGCAG 1048
DB 1297 AGTATCAGAGGATGGGATGATGTCAGTCCACAGAGAAATCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
DB 1357 ACTT 1360
RESULT 15
US-10-843-641A-4854
; Sequence 4854, Application US/10843641A
; Publication No. US2005064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4854
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-4854

Query Match 55.1%; Score 580; DB 22; Length 2500;
Best Local Similarity 75.1%; Pred. No. 4.6e-157;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
QY 89 ATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGCGGGTTTGTCACTTTTCCAAATCCAGTACC 148
DB 397 AGATGAGCGCGTGTGAGAGAAACCCCGCGCGTTTGTCACTTTTCCCAATCCAGTACC 456

QY 149 CTGATATTTGAAAAATGTATAAACAGGCACAGGCTTCTCTCTGACAGCAGAAAGAGTTG 208
DB 457 ATGATATCTGGCAGATGATTAAGAGGCAGAGGCTTCTCTTTGGACCGCCGAGAGGTTG 516
QY 209 ACTTATCAAGAGTCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTACTTCACT 268
DB 517 ACCTCTCCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
QY 269 CTCACATCTTAGCCTTTTTCAGCCAGTGTGAATTTGTAATGAAATTTGGTGGAGC 328
DB 577 CCATGTTCTGGCTTTCTTTGACAGCAAGCGATGTCATAGTAATGAAACCTTGGTGGAGC 636
QY 329 GCTTTAGTCAGGAGTGCAGGTTCCAGAGGCTGCTCTCTTCTATGGCTTTCAAAATTTCA 388
DB 637 GATTTAGCCAAAGTTCAGATTACAGAGCCGCTGTTCTTATGGCTTCCAAATTTGCA 696
QY 389 TCAGAAATGTTCACTCAGAGATGTACAGTTTGTGTAGTACACCTTACATCAGAGATCCA 448
DB 697 TGAAGAACATACATTCGAAATGTATAGTCTTCTTATTTGACACTTACATAAAGATCCA 756
QY 449 AGAAAGGGAATTTTATTAATGCAATGAAACCATGCCCTATGTTTAAAGAAAAGCAG 508
DB 757 AAGAAAGGGAATTTCTCTCAATGCCATGAAACGATGCCCTTGTGTCAAGAAAGAGGCG 816
QY 509 ATTGGGCTTTCGATGAGATAGAGATAGAAAATCTACTTTTGGGAAAGAGTGGTGGCT 568
DB 817 ACTGGGCTTTCGCTGATTTGGGACAAAGAGGCTACTATGGTGAACGTTGTGTAGCT 876
QY 569 TTGCTGCTGTAAGAGGAGTTTCTCTCTCAGGATCTTTTGTGCTATATTTCTGGTAAAGA 628
DB 877 TTGCTGAGTGAAGGAGTTTCTTTTTCGCTTCTTTTGGCTCGATATTTCTGGCTCAAGA 936
QY 629 AGAGAGTCTTATGCGCAGGACTCACTTTTTCATGAACTCATCAGCAGAGATGAAGAC 688
DB 937 AACGAGGACTGATGCCCTGGCTCACAATTTCTAATGAACCTTATTAGCAGAGATGAGGTT 996
QY 689 TTCACTGTGACTTTGCTTGGCTGATGTTCCAAATCTAGTAAATAAGCCTTTCAGAAAGAA 748
DB 997 TACACTGTGATTTTGTCTGCTGATTTCAACACCTGTTACACAACCATCGGAGGAGA 1056
QY 749 GGGTCAGGAGATCAATTTGATGCTGTCAAAATTTGAGCAGGAGTTTAAACAGAGCCT 808
DB 1057 GAGTAAGAGAAATTAATCAATGCTGTTCCGATAGAACAGAGTTCTCTCACTGAGCCT 1116
QY 809 TGCCAGTTGGCCCTCAITGGAAATGCAITTTTGATGAAACAGTACATTTGAGTTTGTAG 868
DB 1117 TGCCGTGTAAGCTCATTTGGGATGAATTCACCTCAATGAAGCAATACATTTGAGTTGTG 1176
QY 869 CTGACAGATTAATTTGTTGGAACCTTGAATTCCTCAAGGTTTTCAGGAGAAAATCCCTTTG 928
DB 1177 CAGACAGACTTATGCTGGAACCTGGGTTTTCAGAGTTTTCAGAGTAGAGAACCCATTG 1236
QY 929 ATTTTATGAAACATTTCTTTTGAAGGAAAAAATAATTTCTTTGAGAAAACGAGTTTCAG 988
DB 1237 ACTTTATGAGAAATATTTCACTGGAAGAAAGACTAACTTTCTTTGAGAGAGTAGGCG 1296
QY 989 AGTATCAGGCTTTTGCACTTATGGCAGAAACACAGATAACGCTTTCACCTTGGATGCA 1048
DB 1297 AGTATCAGAGGATGGGAGTGTGTCAGTCCACAGAGAAATCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
DB 1357 ACTT 1360

Search completed: October 30, 2005, 06:24:12
Job time : 679.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 30, 2005, 03:23:39 ; Search time 62.25 Seconds
(without alignments)
13084.616 Million cell updates/sec

Title: US-10-698-228-12

Perfect score: 1846

Sequence: 1 atggggaccggaaagcc.....tcacctgtgacgatttt 1053

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xlh
-Q=/cgn2.1/USPTO spool/US10698228/runat_26102005_100607_3971/app_query.fasta_1.2446
-DB=A_Geneseq_16Dec04 -QFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228 @CGN_1_1_140 @runat_26102005_100607_3971 -NCPU=6 -ICPU=3
-NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1821	98.6	351	3 AAY84439	Aay84439 Amino aci
2	1821	98.6	351	4 AAB69050	Aab69050 Human rib
3	1817	98.4	351	4 AAB93497	Aab93497 Human pro
4	1454	78.8	389	4 AAB69051	Aab69051 Human R2
5	1454	78.8	389	6 ABU07433	Abu07433 Protein d
6	1454	78.8	389	7 ADJ68753	Adj68753 Human hea
7	1454	78.8	389	8 ADJ66564	Adj66564 Ribonucle
8	1454	78.8	389	8 ADK70435	Adk70435 Respirato
9	1454	78.8	389	8 ADN03789	Adn03789 Antipsori
10	1454	78.8	389	8 ADN04444	Adn04444 Antipsori

11	1454	78.8	389	8 ADO19226	Ado19226 Human PRO
12	1454	78.8	389	8 ADQ09272	Adq09272 Human REM
13	1454	78.8	389	8 ABM80258	Abm80258 Tumour-as
14	1454	78.8	413	3 AAB43902	Aab43902 Human can
15	1454	78.8	453	4 AAU28017	Aau28017 Human con
16	1454	78.8	453	4 ABG15287	Abg15287 Novel hum
17	1429	77.4	390	2 AAW41765	Aaw41765 Human rib
18	1425.5	77.2	386	8 ADO57309	Ado57309 Kidney de
19	1348	73.0	393	4 ABB64130	Abb64130 Drosophil
20	1242	67.3	430	8 ADN21469	Adn21469 Bacteri
21	1230	66.6	405	6 ABJ26187	Abj26187 Aspergill
22	1214	65.8	381	8 ADN23033	Adn23033 Bacteri
23	1192.5	64.6	381	6 ABJ25587	Abj25587 Aspergill
24	1192	64.6	413	5 ABP73700	Abp73700 Candida a
25	1176.5	63.7	321	8 ADS21332	Ads21332 Bacteri
26	1169.5	63.4	399	6 ABR53196	Abr53196 Protein s
27	1169.5	63.4	399	7 ADK63240	Adk63240 Disease t
28	1169.5	63.4	399	8 ADS43740	Ads43740 Bacteri
29	1161.5	62.9	391	8 ADS44135	Ads44135 Bacteri
30	1161.5	62.9	401	7 ADB70101	Adb70101 C. neofo
31	1132	61.3	352	3 AAG05697	Aag05697 Arabidops
32	1130	61.2	329	3 AAG51528	Aag51528 Arabidops
33	1130	61.2	329	3 AAG05699	Aag05699 Arabidops
34	1130	61.2	332	3 AAG51527	Aag51527 Arabidops
35	1130	61.2	332	3 AAG05698	Aag05698 Arabidops
36	1117	60.5	347	3 AAG14546	Aag14546 Arabidops
37	1112.5	60.3	330	3 AAG14548	Aag14548 Arabidops
38	1112.5	60.3	333	3 AAG14547	Aag14547 Arabidops
39	1099	59.5	314	3 AAG51529	Aag51529 Arabidops
40	1091	59.1	341	3 AAG41395	Aag41395 Arabidops
41	1091	59.1	363	3 AAG41394	Aag41394 Arabidops
42	1084.5	58.7	330	3 AAG41396	Aag41396 Arabidops
43	1072.5	58.1	274	2 AAW26418	Aaw26418 Swinepox
44	1072.5	58.1	274	4 AAB68239	Aab68239 Protein e
45	977.5	53.0	322	2 AAR72736	Aar72736 Plasmodiu

ALIGNMENTS

RESULT 1

AAY84439

ID AAY84439 standard; protein; 351 AA.

XX AAY84439;

AC AAY84439;

XX 25-JUL-2000 (first entry)

XX Amino acid sequence of a human RNA-associated protein.

XX Human; RNA-associated protein; cell proliferation; cancer; inflammation;
KW immune response; reproductive disorder; actinic keratosis;
KW atherosclerosis; arteriosclerosis; bursitis; cirrhosis; hepatitis;
KW mixed connective tissue disease; myelofibrosis; primary thrombocythemia;
KW paroxysmal nocturnal hemoglobinuria; polycythemia vera; psoriasis;
KW trauma.

XX Homo sapiens.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 18 /note= "potential phosphorylation site"

FT Modified-site 23 /note= "potential glycosylation site"

FT Modified-site 29 /note= "potential phosphorylation site"

FT Modified-site 38 /note= "potential phosphorylation site"

FT Modified-site 39 /note= "potential phosphorylation site"

FT Modified-site 65 /note= "potential phosphorylation site"

FT Modified-site 87 /note= "potential phosphorylation site"

FT Modified-site 87 /note= "potential phosphorylation site"

FT Modified-site 182 /note= "potential phosphorylation site"
 FT Modified-site 225 /note= "potential phosphorylation site"
 FT Modified-site 314 /note= "potential glycosylation site"
 FT Modified-site 329 /note= "potential phosphorylation site"
 XX
 PN WO200015799-A2.
 XX
 PD 23-MAR-2000.
 XX
 PF 17-SEP-1999; 99WO-US021688.
 XX
 PR 17-SEP-1998; 98US-00156039.
 PR 22-SEP-1998; 98US-00158720.
 PR 04-NOV-1998; 98US-00186815.
 PR 08-APR-1999; 99US-0128660P.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Tang YT, Corley NC, Guegler KJ, Gorgone GA, Patterson C;
 PI Hillman JL, Baughn MR, Lal P, Azimzai Y, Yue H, Yang J;
 XX
 DR WPI: 2000-271437/23.
 DR N-PSDB; AAI12411.
 XX
 PR New polypeptides and polynucleotides, useful for preventing and treating
 PT a disorder associated with increased or decreased expression of RNA
 PT associated proteins.
 XX
 PS Claim 1; Page 103-104; 131pp; English.
 XX
 CC The present sequence represents a human RNA-associated protein. The
 CC expression of RNA-associated proteins is closely associated with
 CC reproductive tissues, nervous tissues, cell proliferation including
 CC cancer, inflammation and immune responses, and so they may be used for
 CC diagnosis, treatment or prevention of cell proliferative, diseases and
 CC immune/inflammatory disorders, and reproductive disorders. Diseases and
 CC disorders which may be treated include actinic keratosis,
 CC atherosclerosis, arteriosclerosis, bursitis, cirrhosis, hepatitis, mixed
 CC connective tissue disease, myelofibrosis, paroxysmal nocturnal
 CC hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia
 CC and cancers, and trauma
 XX
 SQ Sequence 351 AA;
 XX
 Alignment Scores:
 Pred. No.: 6 81e-210 Length: 351
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.65% Indels: 0
 DB: 3 Gaps: 0
 US-10-698-228-12 (1-1053) x AAY84439 (1-351)
 QY 1 ATGGGCGACCCGGAAGCGCGGCGGTGGATCAGGATGAGAGATCATCTTCA 60
 DB 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
 QY 61 GACACCAACGAAGTGAATGAAGTCAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 120
 DB 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
 QY 121 CGGTGTTGTCATCTTCCATCCAGTACCCTGATTTGGAAATGTATAAACAGGCACAG 180
 DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetIleGlnAlaGln 60
 QY 181 GCTTCCTCTTCGACAGACGAAGAGGTGTGACTTATCAAGGATCTCCCTCAGTGAACAAG 240
 DB 61 AlaSerPheTrpThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80

QY 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTAGCCTTTTGTGACGCCAGTGAT 300
 DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAspSerAsp 100
 QY 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGGTGCAGGTCCAGAGGCT 360
 DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 QY 361 CGCTGTTTCTATGGCTTTCAAATCTCATCAGAGATGTTCTCACTCAGAGATCTACAGTTTG 420
 DB 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 QY 421 CTGATAGACACTTACATCAGATGCCCAAGAAAGGCAATTTTATTAAATGCAATTGAA 480
 DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATGGGCCCTTGGATGGATGAGATAGAGAAAA 540
 DB 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTCTGTAGAGGAGTTTCTTCTCAGA 600
 DB 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 QY 601 TCTTTTGTCTGCTATATTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCCTTTTCC 660
 DB 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 QY 661 AATGAACATCATCAGCAGATGAAGGACTTCTCAGTGTGCTGCTGCTGCTGCTGCTGCTCAA 720
 DB 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 QY 721 TACTTAGTAATAAGCCTTTCAGAAAGAGCGTCCAGGAGATCATTTGTGCTGTCAA 780
 DB 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 QY 781 ATTGAGCAGGAGTCTTTTAAACAGAGCCTTGCAGTTCGCTCATTTGGAATCAATTGCATT 840
 DB 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 QY 841 TTGATGAAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
 DB 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 QY 901 AAGTCTTTCAGGCAGAAAATCCTTTTGATTTTATGGAACAAATCTTTTGAAGGAAAA 960
 DB 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 QY 961 ACAAAATTTCTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1020
 DB 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 QY 1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053
 DB 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 RESULT 2
 AAB69050
 ID AAB69050 standard; protein; 351 AA.
 XX
 AC AAB69050;
 XX
 DT 18-APR-2001 (first entry)
 XX
 DE Human ribonucleotide reductase TP53R2H protein sequence SEQ ID NO:1.
 XX
 KW Human; ribonucleotide reductase; cancer; DNA repair; p53.
 XX
 OS Homo sapiens.
 XX
 PN WO200100799-A1.
 XX

PD 04-JAN-2001.
 XX PF 27-JUN-2000; 2000WO-JP004189.
 XX PR 28-JUN-1999; 99JP-00181131.
 XX PR 06-JUL-1999; 99JP-00192391.
 XX PR 21-JAN-2000; 2000JP-0001770.
 XX (TAKE) TAKEDA CHEM IND LTD.
 PA (NAKA/) NAKAMURA Y.
 XX PI Nakamura Y, Arakawa H, Tanaka H;
 XX DR N-PSDB; AAF32438.
 XX DR WPI; 2001-112446/12.
 XX PT Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
 PT diagnosis, treatment and prevention of cancer.
 XX PS Claim 1; Fig 1-3; 102pp; Japanese.
 CC The present sequence represents a human ribonucleotide reductase
 CC designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of
 CC the DNA repair mechanism and its activity is induced by p53. It can be
 CC used for the treatment, prevention and diagnosis of a wide range of
 CC cancers
 XX SQ Sequence 351 AA;
 Alignment Scores:
 Pred. No.: 6.81e-210 Length: 351
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.65% Indels: 0
 DB: 4 Gaps: 0
 US-10-698-228-12 (1-1053) x AAB69050 (1-351)
 QY 1 ATGGCGACCCGGAAAGCGCGGAGCGGCGCTGGATCAGATGAGAGATCATCTTCA 60
 DB 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
 QY 61 GACACCAACGAAGTCAATTAAGTCAATTAAGTCAATTAAGTCAATTAAGTCAAT 120
 DB 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSer 40
 QY 121 CGGTTTGTTCATCTTCCCAATCCAGTACCCTGATATTGGAAATGATATAACAGGCACAG 180
 DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
 QY 181 GCTTCCTCTGACAGCAGAGAGGTGACTATCAAGATCTCCCTCACTGGAAACAAG 240
 DB 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
 QY 241 CTTAAGCAGATGAGAGTACTTCTCTCATCTTACATCTTACATCTTACATCTTACAT 300
 DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
 QY 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
 DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 QY 361 CGCTGTTTCTATGGCTTCCAAATTCATCGAATGTTTCACTCAGAGATGATACAGTTTG 420
 DB 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 QY 421 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
 DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 QY 481 ACCATGCCCTTGTAAAGAAAAAGCAGATTGGCCCTTGGATGGATAGCAGATAGAAA 540

Db 161 ThrMetProTyrValLysLysLysAlaAspTrrAlaLeuArgTrrPilleAlaAspArgLys 180
 QY 541 TCTACTTTTGGGAAAGAGTGGCTTGTCTGTAGAGGAGTCTTCTTCTCAGGA 600
 Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACATTTTCC 660
 Db 201 SerPheAlaAlaIlePheTrrPheLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 QY 661 AATGAATCATCAGCAGAGATGAAGACTTCTCAGTGTGCTTGTGCTGCTGCTGCTGCTCAA 720
 Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 QY 721 TACTTAGTAATAGCTTCCAGAGAAAGGTGAGGAGATCATTTGTGCTGTCTCAA 780
 Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 QY 781 ATTGAGCAGGAGTCTTAAACAGAGAGCTTCCAGTGTGCTGCTGCTGCTGCTGCTGCTT 840
 Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 QY 841 TTGATGAAACAGTACATTTAGTGTGAGCTGAGCAGATTTCTGTGGAACTTGGATTTCA 900
 Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 QY 901 AAGTTTTCAGCAGAGAAATCTTTGATTTTATGGAACATTTCTTTAGAGGAAA 960
 Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 QY 961 ACAATTTCTTTCAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTATGCGCAGAAACC 1020
 Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 QY 1021 ACAGATAACGTTCTTCCCTTGGATGAGATTTT 1053
 Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 RESULT 3
 AAB93497
 ID AAB93497 standard; protein; 351 AA.
 XX AC AAB93497;
 XX DT 26-JUN-2001 (first entry)
 DE Human protein sequence SEQ ID NO:12811.
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX OS Homo sapiens.
 XX PN EP1074617-A2.
 XX PD 07-FEB-2001.
 XX PF 28-JUL-2000; 2000EP-00116126.
 XX PR 29-JUL-1999; 99JP-00248036.
 XX PR 27-AUG-1999; 99JP-00300253.
 XX PR 11-JAN-2000; 2000JP-00118776.
 XX PR 02-MAY-2000; 2000JP-00183767.
 XX PR 09-JUN-2000; 2000JP-00241899.
 XX (HELI-) HELIX RES INST.
 XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or

PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 PS Claim 8; SEQ ID NO 12811; 2537pp + Sequence Listing; English.
 XX
 XX The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 XX
 SQ Sequence 351 AA;

Alignment Scores:

Pred. No.: 351
 Score: 1817.00 Matches: 350
 Percent Similarity: 100.00% Conservatives: 1
 Best Local Similarity: 99.72% Mismatches: 0
 Query Match: 98.43% Indels: 0
 DB: 4 Gaps: 0

US-10-698-228-12 (1-1053) x AAB93497 (1-351)

QY 1 ATGGCGGACCCGGAAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
 DB 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
 QY 61 GACACCAAGCAAGTGAATTAAGTCAATGAGAGCCACTCTAAGAAGAGTTCTCC 120
 DB 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
 QY 121 CGGTTTGTCTATCTTCCATCCAGTACCCTGATATTGGAAATGTATAACAGGCACAG 180
 DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
 QY 181 GCTTCTCTTCTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGACCAAG 240
 DB 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 QY 241 CTTAAGCGATGAGAGAGTACTTCTCTCTACATCTTACGCTTTTTCAGCCAGTGTAT 300
 DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
 QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
 DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 QY 361 CGCTGTTCTATGGCTTTCAAATTTCTATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
 DB 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 QY 421 CTGATAGACATTACATCAGAGATCCCAAGAAAGGAAATTTTATTATGCAATTGAA 480
 DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATGGCCCTTGGATGAGATAGCAAAAA 540

DB 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 QY 541 TCTACTTTTGGGGAAGAGAGTGGCCCTTTCCTCTAGAGAGGATTTCTTCTCAGGA 600
 DB 181 SerThrPheGlyGluArgValAlaAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 QY 601 TCTTTTGTCTCTATATCTCGCTAAAGAGAGAGGTCTTATGCCAGGACTCACATTTTCC 660
 DB 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 QY 661 AATGAATCATCAGCAGAGATGAAGACTTCACGTGACATTTGCTGCTGCTGATGTTCCAA 720
 DB 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 QY 721 TACTTAGTAATTAAGCCTTCAGAGAAGAGGTCCAGGAGATCATTTGATGCTGTCAA 780
 DB 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 QY 781 ATTGAGCAGGAGTTTAAACAGAGCCCTTGCAGTTCGCTTCATTTGGAATGCAATTGCATT 840
 DB 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyValAsnCysIle 280
 QY 841 TTGATGAAACAGTACATTTGATGTTGTAGTCACAGATTACTTGTGGAACCTTGGATTCTCA 900
 DB 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 QY 901 AAGCTTTTTCAGCAGAAAAATCCTTTTGATTTTATGAAAAACATTTCTTTAGAGGAAAA 960
 DB 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 QY 961 ACAAAATTTCTTTCAGAAACAGATTTCAGAGTATCAGCGTTTTCAGATTATGCGCAGAAACC 1020
 DB 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 QY 1021 ACAGATAACCTCTTTCACCTTTGGATGCGAGATTTT 1053
 DB 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 4
 AAB69051
 ID AAB69051 standard; protein; 389 AA.
 XX
 AC AAB69051;
 XX
 DT 18-APR-2001 (first entry)
 XX
 DE Human R2 protein sequence SEQ ID NO:5.
 DE
 KW Human; ribonucleotide reductase; cancer; DNA repair; p53.
 XX
 OS Homo sapiens.
 XX
 PN WO200100799-A1.
 XX
 PD 04-JAN-2001.
 XX
 PF 27-JUN-2000; 2000WO-JP004189.
 XX
 PR 28-JUN-1999; 99JP-00181131.
 PR 06-JUL-1999; 99JP-00192391.
 PR 21-JAN-2000; 2000JP-00017770.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 PA (NAKA/) NAKAMURA Y.
 XX
 PI Nakamura Y, Arakawa H, Tanaka H;
 XX
 XX WPI; 2001-112446/12.
 XX
 XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
 PT diagnosis, treatment and prevention of cancer.

PS Example 2; Page 92-94; 102pp; Japanese.

XX The present invention describes a human ribonucleotide reductase

CC designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of

CC the DNA repair mechanism and its activity is induced by p53. It can be

CC used for the treatment, prevention and diagnosis of a wide range of

CC cancers. The present sequence represents the human R2 protein which is

CC used in an example from the present invention

XX Sequence 389 AA;

XX

Alignment Scores:

Pred. No.:	1,42e-165	Length:	389
Score:	1454.00	Matches:	269
Percent Similarity:	94.08%	Conservative:	33
Best Local Similarity:	83.80%	Mismatches:	19
Query Match:	78.76%	Indels:	0
DB:	4	Gaps:	0

US-10-698-228-12 (1-1053) x AAB69051 (1-389)

QY 91 GAAGAGCCACTCTTAAGAAAGATTCTCGCGGTTGTGTCATCTTCCATCCAGTACCT 150

DB 69 AspGluProLeuLeuArgGluAsnProArgArgPheValIlePheProIleGluTyrHis 88

QY 151 GATATTGGAAATGTATAAGACGACAGGCTTCTTGGACACGAGAGGTTGAC 210

DB 89 AspIleTrpGlnMetTyrLysAlaGluAlaSerPheTrpThrAlaGluGluValAla 108

QY 211 TTATCAAGAGTCTCCCTCACTGGAAACAGCTTAAAGCAGATGAGAGTACTTCACTCT 270

DB 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluLysArgTyrPheIleSer 128

QY 271 CACATCTTAGCTTTTTCAGCCAGTGTGAATTTGTAATGAAATTTGGTGAGCGC 330

DB 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148

QY 331 TTATGTCAGGAGTCCAGTTCAGGAGCTCGCTGTTCTATGGCTTTCAATTCCTC 390

DB 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168

QY 391 GAGATGTTCACTCAGAGATGTACAGTGTGCTGATAGACACTTACATCAGAGATCCCAAG 450

DB 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188

QY 451 AAAAGGGAATTTTATTAATCAATTTGAACCATGCCCTATGTTTAAAGAAAGAGAGAT 510

DB 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208

QY 511 TGGGCTTGGAGTGTAGCAGATAGAGAAATCTACTTTGGGAAAGAGTGGTGCTTT 570

DB 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 228

QY 571 GCTGCTGTAGAAGAGTTTCTTCTCAGGATCTTTGCTGCTATATCTTGGCTAAGAAG 630

DB 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248

QY 631 AGAGTCTTATGCCAGGACTCACTTTTCCATGAATCACTCAGCAGAGATGAAGACTT 690

DB 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268

QY 691 CACTGTGACTTGTGCTGCTGATGTTCCTCAATCTAGTAAATAGCCTTCAGAAAGAG 750

DB 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288

QY 751 GTCAGGAGATCATTTGTTGATCTGTCAAAATGAGCAGAGTTTAAACAGAGCTTG 810

DB 289 ValArgGluIleIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308

QY 811 CCAGTTGGCTCATTTGGAATGATTTGCAATTTGATGAACAGTACATTCAGTTTGTAGCT 870

DB 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328

QY 871 GACAGATTACTGTGGAACCTTGATTTCTCAAGGTTTTCAGGCAGAAAAATCCTTTGAT 930

DB 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348

QY 931 TTATGGAACACATTTCTTTAGAGGAAAAACAATTTCTTTGAGAAACGAGTTTCAGAG 990

DB 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368

QY 991 TATCAGCGTTTTCAGTTATGGCAGAAACACAGATACGCTTCTCACCTTGGATCGAGAT 1050

DB 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388

QY 1051 TTT 1053

DB 389 Phe 389

RESULT 5

ABU07433

ID ABU07433 standard; protein; 389 AA.

XX AC ABU07433;

XX 28-JAN-2003 (first entry)

XX Protein differentially regulated in prostate cancer #36.

XX Prostate cancer; gene expression; differential regulation;

KW molecular marker; drug target; cancer detection; cancer diagnosis;

KW cancer staging; cancer grading; cancer assessing; cancer monitoring.

XX Homo sapiens.

XX W0200281638-A2.

XX 17-OCT-2002.

XX 08-APR-2002; 2002WO-US010824.

XX 06-APR-2001; 2001US-0281731P.

PR 06-APR-2001; 2001US-0281732P.

XX (ORIG-) ORIGENE TECHNOLOGIES INC.

PI Sun 2, Jay G;

XX WPI; 2003-058520/05.

DR N-PSDB; ABX10335.

XX Novel genes which are differentially regulated in prostate cancer, useful

PT for diagnosing prostate cancer in prostate tissue sample and assessing

PT therapeutic or preventive intervention in prostate cancer patients.

XX Claim 1; Page 273-274; 416pp; English.

XX The invention describes genes (I) which are differentially regulated in

CC prostate cancer. (I) is useful for diagnosing a prostate cancer in a

CC sample comprising prostate tissue, which involves determining the number

CC of target genes which are differentially-regulated in the sample, where

CC the number is indicative of the probability that the sample comprises

CC prostate cancer. (I) is useful for assessing a therapeutic or preventive

CC intervention in a subject having a prostate cancer, which involves

CC determining the expression levels in a sample comprising prostate tissue

CC of target genes which are differentially-regulated in prostate cancer.

CC Preferably, the expression levels of at least 10 genes are determined.

CC (I) is also useful for identifying agents that modulate a biological

CC activity of a polypeptide differentially-regulated in prostate cancer

CC cells, which involves contacting a polypeptide differentially-regulated

CC in prostate cancer cells with a test agent under conditions effective for

CC the test agent to modulate a biological activity of the polypeptide, and

CC determining whether the test agent modulates the biological activity. (I)

CC is useful as molecular markers, as drug targets, and for detecting,

CC diagnosing, staging, grading, assessing, monitoring, prognosticating,

CC preventing or treating, determining predisposition to diseases and

CC conditions especially relating to prostate cancer. (I) and its expression
CC products are used in the diagnostic test to assay for presence of cancer
CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
CC blood etc. (I) is useful for assessing cancer e.g., to determine the type
CC of cancer, its stage of development, the nature of genetic defect, etc.
CC The polypeptide encoded by (I) can be used as target for therapy or drug
CC discovery. (I) can also be used for expressing the polypeptide and thus
CC for searching specific binding partners of the polypeptide. (I) is useful
CC in therapeutic applications to treat prostate cancer. The identification
CC of specific genes, and groups of genes, expressed in pathways
CC physiologically relevant to prostate cancer permits the definition of
CC functional and disease pathways and the delineation of targets in these
CC pathways which are useful in diagnostic, therapeutic, and clinical
CC applications. This is the amino acid sequence of a protein differentially
CC regulated in prostate cancer

XX
SQ Sequence 389 AA;

Alignment Scores:
Pred. No.: 1,426-165 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservatives: 33
Best Local Similarity: 83.88% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 6 Gaps: 0

US-10-698-228-12 (1-1053) x ABU07433 (1-389)

QY 91 GAAGAGCACTCTAAGAAAGATGCTCGCGGTTGTGCATCTTCCAAATCAGTACCTT 150
DB 69 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 88
QY 151 GATATTTGGAAATCTATAACAGGCACAGGCTCTCTCGACAGCAAGAGTTGAC 210
DB 89 AspIleTrpGlnMetTyrIleValAlaGluAlaSerPheTrpThrAlaGluValAla 108
QY 211 TTATCAAGGATCTCCCTCACTGGAACAGCTTAAAGCAGATGAGAAGTACTTCTCT 270
DB 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTTAGCTTTTGGACCGAGTGAATGGAATGTAATGAAATTTGGTGAGCGC 330
DB 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY 331 TTATGTCAGGAGTCCAGGCTCCAGAGCTGCTGTTCTATGCTTTCATCTTCATC 390
DB 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGATGTTCACTCAGAGATGACAGTTGCTGATAGACACTTACATCAGAGATCCCAAG 450
DB 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAAGGGAATTTTATTAATGCAATGGAACCATGCCCTATGTTAAGAAAAAAGCAGAT 510
DB 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
QY 511 TGGGCTTGCAGATGATAGACATGAAATCTACTTTGGGAAAGATGTGGCTTT 570
DB 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaAlaPhe 228
QY 571 GCTGCTGTAGAGGAGTTTCTCTCAGGATCTTTTGTGCTATATCTTGGCTAAAGAG 630
DB 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGTCTTATGCCAGGACTCACTTTTCCATGAACTCATCAGCAGAGATGAAGGACTT 690
DB 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY 691 CACTGTGATCTTGTGCTGATGTTTCCAAATCTAGTAAATAGCCTTTCAGAGAAAGG 750
DB 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGGAGATCATGTTGATGCTGCTCAAAATGAGCAGGAGTTTAAACAGAGCCTTG 810

DB 289 ValArgGluIleIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCTCATTGGAATGCAATTCATTTGATGAACAGTACATTGAGTTTGTAGCT 870
DB 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
QY 871 GACAGATTACTTCTGGAACCTTGATTTCTCAAGGTTTTCAGGCAGAAAAATCCTTTGAT 930
DB 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY 931 TTATCGAAAAACATTTCTTTAGAGGAAAAACAAATTTCTTTTGAGAAACGAGTTTCAGAG 990
DB 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY 991 TATCAGGCTTTTCAGTTATGCGAGAAACACAGATAACGCTTCTCACCTTCGATGCAGAT 1050
DB 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY 1051 TTT 1053
DB 389 Phe 389
RESULT 6
ADJ68753
ID ADJ68753 standard; protein; 389 AA.
AC ADJ68753;
XX
DT 06-MAY-2004 (first entry)
DE Human heat mitochondrial protein as a therapeutic target SeqID559.
XX
KW mitochondrial; human; screening assay; diabetes mellitus;
KW Huntington's disease; osteoarthritis;
KW Leber's hereditary optic neuropathy; LHON;
KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
KW osteopathic; ophthalmological; cytostatic.
XX
OS Homo sapiens.
PN WO2003087768-A2.
XX
PD 23-OCT-2003.
XX
PF 04-APR-2003; 2003WO-US010870.
XX
PR 12-APR-2002; 2002US-0372843P.
PR 17-JUN-2002; 2002US-0389987P.
PR 20-SEP-2002; 2002US-0412418P.
XX
PA (MITO-) MITOKOR.
PA (BUCK-) BUCK INST AGE RES.
XX
PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
PI Warnock DE;
XX
DR WPI; 2003-845369/78.
XX
PT Identifying a mitochondrial target for drug screening assays and for
PT treating diseases associated with altered mitochondrial function,
PT comprises detecting a modified polypeptide in a sample and correlating
PT with the disease.
XX
PS Claim 1; SEQ ID NO 559; 180pp; English.
XX
CC This invention relates to novel mitochondrial targets that can be used
CC for therapeutic intervention in treating a disease associated with
CC altered mitochondrial function. Specifically, it refers to a method for
CC identifying proteins of the human heart mitochondrial proteome that are
CC useful for drug screening assays, as well as therapeutic targets. The

present invention describes a method for identifying such proteins that can be used in the treatment of various diseases associated with altered mitochondrial function including diabetes mellitus, Huntington's disease, osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy ragged red fibre syndrome (MERRF) or cancer. Accordingly, these compositions have neuroprotective, neurotropic, antidiabetic, anticonvulsant, antiarthritic, osteopathic, ophthalmological and cytoskeletal activities. This polypeptide sequence is a human heart mitochondrial protein of the invention.

US-10-698-228-12 (1-1053) X ADJ68753 (1-389)

QY	91	GAAGAGCCATCTCTTAAGAAAGAGTTC	TGCGCGGTTTGTCATCTTTTCCAAATCCAGTACCCCT	150
Db	69	AspGluProLeuLeuArgGluAsnProArg	ArgPheValIlePheProIleGluTyrHis	88
QY	151	GATATTTGGAAAAATGTATAAACAGCA	CAGCCTTCCTTCGGACACGACGAAAGGTTGAC	210
Db	89	AspIleIleTrpGlnMetTyrLysAlaGlu	AlaSerPheTrpThrAlaGluGluValAsp	108
QY	211	TTATCAAAAGGATCTCCCTC	ACTCGGAACAAGCTTAAACGACATGAGAAGTACTTCACTCTCT	270
Db	109	LeuSerLysAspIleIleHIS	TrpGluSerLeuLysProGluGluArgTyrPheIleSer	128
QY	271	CACATCTTAGCCCTTTTTCGACCCAGTC	ATGGAATCTGAAATGAAAAATTTGGTGAGCGC	330
Db	129	HisValLeuAlaPhePheAlaAlaSer	AspGlyIleValAsnGluAsnLeuValGluArg	148
QY	331	TTTATGTCAGGAGGTGCAGGTTCCAGAG	CTCGCTGTGTTCTATGGCTTTCAAATTTCTCATC	390
Db	149	PheSerGlnGluValGlnIleThrGluAla	ArgCysPheTyrGlyPheGlnIleAlaMet	168
QY	391	GAGATGTTCTAC	TACAGATGTACAGTTTGCTGTGATGACACTTACATCAGAGATCCCAAG	450
Db	169	GluAsnIleHisSerGluMetTyrSerLeu	LeuIleAspThrTyrIleLysAspProLys	188
QY	451	AAAAAGGAATTTTATTTAATGCAATTA	GCAATTAAGAACCATGCCCTATGTTAAGAAAAAGCAGAT	510
Db	189	GluArgGluPheLeuPheAsnAlaIleGlu	ThrMetProCysValLysLysLysAlaAsp	208
QY	511	TGGGCTCTGCCATGATGATAGCAGATAGA	AAAAATCTACTTTGGGAAAGAGATGGTGCGCTTT	570
Db	209	TrpAlaLeuArgTrpIleGlyAspLysGlu	AlaThrTyrGlyGluArgValValAlaPhe	228
QY	571	GCTGCTGTAGAAGGAGTTTCTCTCAGGAT	CTTTTCTGCTCTATATTCGGCTAAAGAAG	630
Db	229	AlaAlaValGluGlyIlePhePheSerGlySer	PheAlaSerIlePheTrpLeuLysLys	248
QY	631	AGAGTCTTATGCCAGGACTC	ACTTTTTCCATGAACATCATCAGCAGAGATGAAGACTT	690
Db	249	ArgGlyLeuMetProGlyLeuThrPheSer	AsnGluLeuIleSerArgAspGluGlyLeu	268
QY	691	CACTGTGACTTGTGCTGCTGATGTC	CAATCTAGTAAATAGCCTTCAGAGAAGG	750
Db	269	HisCysAspPheAlaCysLeuMetPhe	LysHisLeuValHisLysProSerGluGluArg	288
QY	751	GTCCGGAGATCATTTGTTGATGCTG	CTCAAAATTCAGCAGAGTTTTTAAACAGAGCCTTG	810
Db	289	ValArgGluIleIleIleAsnAlaValArg	IleGluGlnGluPheLeuThrGluAlaLeu	308
QY	811	CCAGTTGCGCCTCATTTGGAATGAAT	TGCAATTTGATGAAACAGTACATTTGAGTTGTAGCT	870

CC EDTA, and 0.1% SDS for 1.5 hours at 60 deg C. The composition and methods
 CC are useful in diagnosing or treating diseases and disorders, preferably
 CC neurodegenerative diseases. These may also be used as a drug target or in
 CC manufacturing a medicament for the treatment or prevention of the above-
 CC mentioned diseases or disorders. The composition may also be used for
 CC treating cancer. This sequence represents one of the proteins of the
 CC complex of the invention.

XX SQ Sequence 389 AA;

Alignment Scores:
 Pred. No.: 1.42e-165 Length: 389
 Score: 1454.00 Matches: 269
 Percent Similarity: 94.08% Conservative: 33
 Best Local Similarity: 83.80% Mismatches: 19
 Query Match: 78.76% Indels: 0
 DB: 8 Gaps: 0

US-10-698-228-12 (1-1053) x ADJ66564 (1-389)

QY 91 GAAGAGCCACTCTAAGAAGAGTTCTCGCGGTTGTGTCATCTTCCCAATCCAGTACCCT 150
 Db 69 AspGluProLeuLeuArgGluLeuProArgGluPheValIlePheProIleGluTyrHis 88
 QY 151 GATATTGGAAATGTATAACAGGACAGGCTTCTCTGGACAGAGAGAGTTGAC 210
 Db 89 AspIleTrpGlnMetTyrIleLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
 QY 211 TTATCAAGGATCTCCCTCACTGGAAACAAGCTTAAGACAGAGTACTTCATCTCT 270
 Db 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
 QY 271 CACATCTTAGCTTTTTCAGCCAGCTGATGGAATGTAATGAAATTTGGTGGAGCCG 330
 Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
 QY 331 TTATGTCAGAGTGCAGGTTCAGAGGCTCGCTGTTCTATGGCTTCAAAATTTCTATC 390
 Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
 QY 391 GAGATGTTCACTCAGAGATGTACAGTTTCTGATAGACACTTACATCAGAGATCCCAAG 450
 Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
 QY 451 AAAAGGGAATTTTATTAATGCAATTAAGAACCATGCCCTATGTTAAAGAAAAGCAGAT 510
 Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValIleLysLysAlaAsp 208
 QY 511 TGGGCTTGGATGAGTACAGATAGAAATCTACTTTTGGGAAAGAGTGGTGGCTTT 570
 Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaIlePhe 228
 QY 571 GCTGCTGTAGAAGGATTTCTTCTCAGGATCTTTTGTGCTATATTGCTGCTTAAGAAG 630
 Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
 QY 631 AGAGTCTTATGCCAGGACTCACTTTTCCATGAATCACTACGACAGATGAAGACTT 690
 Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
 QY 691 CACTGTGACTTTGCTGCTGCTGATGTTCCCAATCTAGTAAATAGCCTTCAGAGAAAGG 750
 Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
 QY 751 GTCAGGAGATCATTTGATGCTGTCAAAATAGCAGGAGTTTAAACAGAGGCTTG 810
 Db 289 ValArgGluIleIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
 QY 811 CCAGTTGGCTCATTTGGAATGAATTCATTTTGATGAAACAGTACATTCAGTTTGTAGCT 870
 Db 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
 QY 871 GACAGATTACTTGTGGAACCTTGATTTCTCAAGGTTTTCAGGACAGAAATCTTTTGAT 930

Db 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
 QY 931 TTATCGAAACATTTCTTTAGAGGAGAAACAAATTTCTTGAGAAACGAGTTTCAGAG 990
 Db 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPheGluLysArgValGlyGlu 368
 QY 991 TATCAGCGTTTTCAGTTATGCGCAAAACCCACAGATAACCTTTCACCTTCGATGCAGAT 1050
 Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
 QY 1051 TTT 1053
 Db 389 Phe 389
 RESULT 8
 ID ADK70435
 ID ADK70435 standard; protein; 389 AA.
 AC ADK70435;
 DT 06-MAY-2004 (first entry)
 XX Respiratory disease differentially expressed protein #1.
 DE cytosolic; respiratory; antiasthmatic; Gene therapy;
 KW differential gene expression; respiratory disorder; lung cancer;
 KW chronic obstructive pulmonary disease; emphysema; asthma.
 OS Homo sapiens.
 PN WO2003101283-A2.
 PD 11-DEC-2003.
 PF 02-JUN-2003; 2003WO-US017409.
 PR 04-JUN-2002; 2002US-0386005P.
 XX (INCY-) INCYTE CORP.
 PA Rickert PK, Krasnow R;
 DR WPI; 2004-042945/04.
 XX New combination comprising cDNAs and proteins that are differentially
 PT expressed in respiratory disorders, useful for diagnosing or treating
 PT respiratory diseases e.g. lung cancer, chronic obstructive pulmonary
 PT diseases or asthma.
 XX Claim 14; SEQ ID NO 171; 343pp; English.
 CC The invention relates to cDNA sequences that are differentially expressed
 CC in respiratory disorders or their complements or encoded proteins. The
 CC cDNAs and proteins are useful for diagnosing, treating or monitoring
 CC treatment of a subject with a respiratory disease including lung cancer,
 CC chronic obstructive pulmonary diseases, emphysema or asthma. The protein
 CC is also useful for screening molecules or compounds to identify at least
 CC one ligand which specifically binds the protein. It is also useful for
 CC preparing and purifying a polyclonal or monoclonal antibody. This
 CC sequence corresponds to a protein of the invention.
 XX Sequence 389 AA;
 SQ Alignment Scores:
 Pred. No.: 1.42e-165 Length: 389
 Score: 1454.00 Matches: 269
 Percent Similarity: 94.08% Conservative: 33
 Best Local Similarity: 83.80% Mismatches: 19
 Query Match: 78.76% Indels: 0
 DB: 8 Gaps: 0
 US-10-698-228-12 (1-1053) x ADK70435 (1-389)

```
QY 91 GAAGGCCACTCTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTTCCAATCCAGTACCCT 150
DB 92 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 88
QY 151 GATATTGGAAATGTATATAACAGGCACAGGCTTCTTCGACACAGAGAGGTTGAC 210
DB 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
QY 211 TTATCAAGGATCTCCCTCACTCGAACAAGCTTAAAGCAGATCAGAGTACTCTCT 270
DB 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTAGCCTTTTTCAGCCAGTGTAAATGTAATGAAATTTGGTGGAGCGC 330
DB 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY 331 TTATGTCAGAGGTGCAGGTTCCAGAGGCTCGCTTCTATGGCTTCAAAATTCATC 390
DB 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCAG 450
DB 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAAGGGAATTTTATTAATGCAATTTGAAACCATCGCTATCTTAAAGAAAAGCAGAT 510
DB 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
QY 511 TGGGCTTCGAGTGGATAGACATAGAAAATCTACTTTTGGGAAAAGAGTGGGCTTT 570
DB 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValValAlaPhe 228
QY 571 GCTGCTGTAGAGGAGTTTCTTCAGAGATCTTTGCTGCTATATCTGGCTTAAAGAG 630
DB 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGGCTTATCCAGGATCACTTTTCCNATGAATCTCATCAGCAGATGAGACATT 690
DB 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY 691 CACTGTGACTTTGCTGCTGATGTTCCTAATCTTAAATAGCTTTCAGAGAAAGG 750
DB 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGAGATCATTTGTCATGCTGCTCAAAATTCAGCAGAGTTTAAACAGAGCCTTG 810
DB 289 ValArgGluIleIleAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCCTCATTTGGAATGAATTCATTTTCATGAACAGTACATTCAGTTTGAGCT 870
DB 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
QY 871 GACAGATTACTTGTGGAATCTCAAGGTTTTCAGGCAAAATCTCTTTGAT 930
DB 329 AspArgLeuMetCysGluGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY 931 TTATGGAAAAATTTCTTTAGAGGAAAAAATAATTTCTTTGAGAAAAAGAGTTTCAGAG 990
DB 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY 991 TATCAGGTTTGGAGTTATGGCAGAAACCAACAGATAACGCTTTCACCTTGGATGAGAT 1050
DB 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY 1051 TTT 1053
DB 389 Phe 389
RESULT 9
ADN03789
ID ADN03789 standard; protein; 389 AA.
XX
```

```
AC ADN03789;
XX 01-JUL-2004 (first entry)
XX Antipsoriatic protein sequence #91.
XX antipsoriatic; gene therapy; psoriasis; diagnosis.
XX Homo sapiens.
XX W02004028479-A2.
XX 08-APR-2004.
XX 25-SEP-2003; 2003WO-US030907.
XX 25-SEP-2002; 2002US-0414006P.
XX (GETH ) GENENTECH INC.
XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
XX Wu TD;
XX WPI; 2004-305105/28.
XX N-PSDB; ADN03788.
XX New PRO nucleic acid or polypeptide, useful for preparing a
XX pharmaceutical composition for diagnosing or treating psoriasis in a
XX mammal.
XX Claim 9; SEQ ID NO 183; 3069pp; English.
XX The invention relates to novel polynucleotide and polypeptides for
XX treating psoriasis or a sequence having at least 80% identity to the
XX above sequences. The nucleic acid is useful for preparing a composition
XX for diagnosing or treating psoriasis in a mammal. This sequence
XX corresponds to one of the polypeptides of the invention.
XX Sequence 389 AA;
SQ Alignment Scores:
Pred. No.: 1,42e-165 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservations: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 8 Gaps: 0
US-10-698-228-12 (1-1053) x ADN03789 (1-389)
QY 91 GAAGGCCACTCTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTTCCAATCCAGTACCCT 150
DB 69 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 88
QY 151 GATATTGGAAATGTATATAACAGGCACAGGCTTCTTCGACACAGAGAGGTTGAC 210
DB 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
QY 211 TTATCAAGGATCTCCCTCACTCGAACAAGCTTAAAGCAGATCAGAGTACTCTCTCT 270
DB 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTAGCCTTTTTCAGCCAGTGTAAATGTAATGAAATTTGGTGGAGCGC 330
DB 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY 331 TTATGTCAGAGGTGCAGGTTCCAGAGGCTCGCTTCTATGGCTTCAAAATTCATC 390
DB 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCAG 450
DB 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
```

```
QY 451 AAAAGGGAATTTTATTATTAATCAATGAAACCATGCCCTATGTTAAGAAAAAGCAGAT 510
DB 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
QY 511 TGGGCTTCGATGGATAGCAGATAGAAAATCTACTTTTGGGAAAGAGTGGGCTTT 570
DB 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 228
QY 571 GCTGCTGTAGAGGAGTTTCTCTCAGGATCTTTGCTGCTATATCTGGCTAAGAAG 630
DB 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLys 248
QY 631 AGAGGTCTTATGCAGGACTCACTTTTCCAAATGAATCATCAGCAGAGATGAAGACTT 690
DB 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY 691 CACTGTGACTTTGCTTGGCTGATGTTCCAAATCTTAGTAATAAGCCTTCAGAAAGG 750
DB 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 CTCAGGAGATCATTTGATGCTGTCAAAATGACAGAGTTTAAACAGAGCCTTG 810
DB 289 ValArgGluIleIleLeuAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCTCATTTGGAATGAATTCATTTGATGAAACAGTACATTTGAGTTGTAGCT 870
DB 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
QY 871 GACAGATTACTTGTGAACTTGATCTCAAAAGTTTTCAGCAGAAAAATCCTTTGAT 930
DB 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY 931 TTTATGGAACAATTTCTTAGAGAAAAACAAATTTCTTCAGAAAAACGAGTTTCAGAG 990
DB 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY 991 TATCAGCGTTTTCAGTTTGGCAGAAACACAGATAACGCTTTCACCTTGGATGCAGAT 1050
DB 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY 1051 TTT 1053
DB 389 Phe 389
RESULT 10
ADN04444
ID ADN04444 standard; protein; 389 AA.
XX AC ADN04444;
XX DT 01-JUL-2004 (first entry)
XX DE Antipsoriatic protein sequence #415.
XX KW antipsoriatic; gene therapy; psoriasis; diagnosis.
XX OS Homo sapiens.
XX PN W02004028479-A2.
XX PD 08-APR-2004.
XX PF 25-SEP-2003; 2003WO-US030907.
XX PR 25-SEP-2002; 2002US-0414006P.
XX PA (GETH ) GENENTECH INC.
XX PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WJ,
XX WU Wu TD;
XX WPI; 2004-305105/28.
```

```
DR N-PSDB; ADN04443.
XX New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.
XX Claim 9; SEQ ID NO 838; 3069pp; English.
XX The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polypeptides of the invention.
XX Sequence 389 AA;
SQ
Alignment Scores:
Pred. No.: 1,42e-165 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 8 Gaps: 0
US-10-698-228-12 (1-1053) x ADN04444 (1-389)
QY 91 GAAGGCCACTCTTAAGAAAGAGTTCTCGCGGTTTCTCATCTTCCAAATCCAGTACCCT 150
DB 69 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluThrHis 88
QY 151 GATATTGGAAAAATGTATAAACAGGCAGCGCTTCTCTGGACAGCAGAGAGAGTTGAC 210
DB 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluValAsp 108
QY 211 TTATCAAGAGATCCCTCTCACTGGAACAAGCTTAAGCAGATGAGAGTACTTCATCTCT 270
DB 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTTAGCCTTTTTCAGCCAGTGTAGTAAATGAAATTTGCTGGAGGC 330
DB 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY 331 TTTAGTCAGGAGGTGCGAGTTCAGAGGCTCGCTGTTTCTATGGCTTTCAAAATTTCTCATC 390
DB 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGAATGTTCACTCAGAGATGTACAGTTTCTGTAGACACTTATCATCAGAGATCCCCAAG 450
DB 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAAGGGAATTTTATTATTAATGCAATTTGAAACCATGCCCTATGTTAAGAAAAAGCAGAT 510
DB 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
QY 511 TGGGCTTCGATGGATAGCAGATAGAAAATCTACTTTTGGGAAAGAGTGGGCTTT 570
DB 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 228
QY 571 GCTGCTGTAGAGGAGTTTCTCTCAGGATCTTTTCTGCTATATCTGGCTAAGAAG 630
DB 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLys 248
QY 631 AGAGGTCTTATGCAGGACTCACTTTTCCAAATGAATCATCAGCAGAGATGAAGACTT 690
DB 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY 691 CACTGTGACTTTGCTTGGCTGATGTTCCAAATCTTAGTAATAAGCCTTCAGAAAGG 750
DB 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGAGATCATTTGATGCTGTCAAAATTCAGCAGAGTTTAAACAGAGCCTTG 810
DB 289 ValArgGluIleIleLeuAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
```

New PRO nucleic acid or polypeptide, useful for preparing a pharmaceutical composition for diagnosing or treating psoriasis in a mammal.

Claim 9; SEQ ID NO 838; 3069pp; English.

The invention relates to novel polynucleotide and polypeptides for treating psoriasis or a sequence having at least 80% identity to the above sequences. The nucleic acid is useful for preparing a composition for diagnosing or treating psoriasis in a mammal. This sequence corresponds to one of the polypeptides of the invention.

Sequence 389 AA;

Alignment Scores:

Pred. No.: 1,42e-165 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 8 Gaps: 0

US-10-698-228-12 (1-1053) x ADN04444 (1-389)

QY 931 TTTATGGAACATTTCTTTAGAGAAACAAATTTCTTTGAGAAACGAGTTTCAGAG 990
 Db PheMetGluAsnIleSerLeuGluGlyLysThrAsnPheGluLysArgValGlyGlu 368
 QY 991 TATCAGCGTTTGGAGTTATGTCAGAAACACAGATAACGTTTACCTTGATGAGAT 1050
 Db 369 TyrGlnArgMetGlyValMetSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
 QY 1051 TTT 1053
 Db 389 Phe 389
 RESULT 12
 ID ADQ09272 standard; protein; 389 AA.
 XX ADQ09272;
 AC ADQ09272;
 XX 23-SEP-2004 (first entry)
 DT Human RRM2 protein SEQ ID NO:457.
 DE
 XX thanatos-associated protein; THAP; THAP responsive gene; THAP family;
 KW THAP responsive element; angiogenesis; inflammation; metastasis; cancer;
 KW apoptosis; cardiovascular disease; neurodegenerative disease; chemokine;
 KW antiangiogenic; antiinflammatory; cardiovascular; cytostatic;
 KW neuroprotective; osteopathic; THAP modulator; THAP synthesis modulator;
 KW human.
 XX
 OS Homo sapiens.
 PN WO2004055050-A2.
 XX
 PD 01-JUL-2004.
 XX
 PF 10-DEC-2003; 2003WO-IB006434.
 PR 10-DEC-2002; 2002US-0432699P.
 PR 03-JUL-2003; 2003US-0485027P.
 XX
 PA (ENDO-) ENDOCUBE SAS.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 PI Girard J, Amalric F, Roussigne M, Clouaire T;
 XX N-PSDB; ADQ09273.
 DR WPI; 2004-525034/50.
 DR N-PSDB; ADQ09273.
 XX
 PT Modulating expression of a Thanatos (death)-Associated Protein (THAP)
 PT responsive gene for preventing or treating e.g. cancer or inflammation,
 PT comprises modulating the interaction of a THAP polypeptide with a nucleic
 PT acid.
 XX
 PS Example 47; SEQ ID NO 457; 612pp; English.
 XX
 CC The present invention describes a method for modulating the expression of
 CC a thanatos (death)-associated protein (THAP) responsive gene. The method
 CC comprises modulating the interaction of a THAP-family polypeptide or its
 CC biological fragment with a nucleic acid, and so enhancing or repressing
 CC the expression of the THAP responsive gene. Also described: (1) a method
 CC of modulating the expression of a gene responsive to a THAP/chemokine
 CC complex; (2) a pharmaceutical composition comprising a THAP responsive
 CC element in a pharmaceutical carrier; (3) a transcription factor decoy
 CC consisting essentially of a THAP responsive element; (4) a cell
 CC comprising a transcription factor decoy described above; (5) methods of
 CC modulating the interaction between a nucleic acid and a THAP-family
 CC polypeptide or its biological fragment, or a nucleic acid and a
 CC THAP/chemokine complex; (6) a vector packaging cell line comprising a
 CC cell comprising a viral vector which comprises a promoter operably linked
 CC to a nucleic acid encoding a THAP-family polypeptide or its biological
 CC fragment; (7) a method of constructing a cell which expresses a
 CC recombinant THAP-family polypeptide; (8) a method of ameliorating
 CC symptoms associated with a condition mediated by a THAP/chemokine complex

CC ; (9) methods of identifying a test compound that modulates transcription
 CC at a THAP responsive element or that modulates the transport of a
 CC chemokine into the nucleus; (10) methods for reducing the symptoms
 CC associated with a condition selected from excessive or insufficient
 CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
 CC or insufficient apoptosis, cardiovascular disease and neurodegenerative
 CC diseases; symptoms associated with a condition resulting from the
 CC activity of a chemokine or a THAP-family polypeptide in an individual; or
 CC symptoms associated with transcriptional repression or activation
 CC mediated by a THAP-family polypeptide in an individual; (11) a vector
 CC comprising a THAP responsive promoter operably linked to a nucleic acid
 CC encoding a detectable product; (12) a genetically engineered cell
 CC comprising the vector described above or that expresses a THAP-family
 CC polypeptide or its biological fragment; (13) an in vitro transcription
 CC reaction comprising a nucleic acid comprising a THAP responsive promoter,
 CC ribonucleotides and an RNA polymerase; and (14) an isolated mutant THAP-
 CC family polypeptide that does not bind to a chemokine. The pharmaceutical
 CC composition has antiangiogenic, antiinflammatory, cardiovascular,
 CC cytostatic, neuroprotective and osteopathic activities, and can be used
 CC as a THAP and THAP synthesis modulator. The composition can be used for
 CC modulating the expression of a THAP responsive gene. Modulation is useful
 CC for reducing symptoms of conditions such as excessive or insufficient
 CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
 CC or insufficient apoptosis, cardiovascular disease or neurodegenerative
 CC diseases. The present sequence is used in the exemplification of the
 CC present invention.
 XX
 SQ Sequence 389 AA;

Alignment Scores:

Pred. No.:	1,42e-165	Length:	389
Score:	1454.00	Matches:	269
Percent Similarity:	94.08%	Conservative:	33
Best Local Similarity:	83.80%	Mismatches:	19
Query Match:	8	Indels:	0
DB:		Gaps:	0

US-10-698-228-12 (1-1053) x ADQ09272 (1-389)

QY	91	GAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTTCCAAATCAGTACCT	150
Db	69	AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis	88
QY	151	GATATTGGAAATGTATAACAGGCACAGCTTCCTTCGACAGCAGAGAGGTTCCAC	210
Db	89	AspIletrpGlnMetTyrLysLysAlaGluAlaSerPheItrpAlaGluGluValAsp	108
QY	211	TTATCAAGAGGATCTCCCTCACTCGAACCAAGCTTAAAGCAGATGAGAGTACTTCATCT	270
Db	109	LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer	128
QY	271	CACATCTTAGCCCTTTTTCGACGAGTGATGGAATTTAAATGAAATTTGGTGGAGCGC	330
Db	129	HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg	148
QY	331	TTTAGTCAGAGGTGTCAGGTTCCAGAGCTCGCTTCTATGGCTTTTCAATTCATCTCATC	390
Db	149	PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet	168
QY	391	GAGAAATGTTCCACTCAGAGATGTACAGTTTGCTGTATAGACCTTACATCAGAGATCCCAAG	450
Db	169	GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys	188
QY	451	AAAAGGGAATTTTATTATTAATGCAATTTGAAACCATGCCCTATGTTAAGAAAAACAGAT	510
Db	189	GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAsp	208
QY	511	TGGGCTTGGGATGATAGCAGATAGAAAATCTACTTTTGGGGAAAGAGTGGTGCCTTT	570
Db	209	TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe	228
QY	571	GCTGCTGTACAAGAGGTTTCTTCTCAGGATCTTTTGTGCTGTATATCTTCTGCTAAAGAG	630

Db 229 AlaAlaValGluGlyLeuPhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGGCTTATGCGCAGGACTCACTTTTCCAACTGAACTCATCAGCAGAGATGAAGGACTT 690
Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY 691 CACTGTGACTTCTGCTGCTGATGTTCCAACTATTAAGTAAAGCCTTCAGAAAGG 750
Db 269 HisCysAspPheAlaCysLeuMetPheLeuHisLeuValHisLysProSerGluGluArg 288
QY 751 CTCAGGAGATCATGTTGATGCTGTCGCAAAATTGAGCAGGAGTTTAAACAGAGCCCTG 810
Db 289 ValArgGluIleIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCCTCATTCGAATGAATTGCAATTTTCATGAACAGTACATTCAGTTTGTAGCT 870
Db 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTrpIleGluPheValAla 328
QY 871 GACAGATTACTTGTGAACCTTGGATTCTCAAGGTTTTCAGGCAGAAAATCCTTTTGAT 930
Db 329 AspArgLeuMetLeuGluGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY 931 TTTATGGAACAACTTCTTAGAAGGAGAAACAAATTTCTTCAGAAACGAGTTTCAGAG 990
Db 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY 991 TATCAGCGTTTTCAGTATGCGCAGAAACACAGATAACGCTCTCACCCTTGGATGCAGAT 1050
Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY 1051 TTT 1053
Db 389 Phe 389

RESULT 13

ABM80258
ID ABM80258 standard; protein; 389 AA.

XX AC ABM80258;
XX DT 18-NOV-2004 (first entry)
XX DE Tumour-associated antigenic target (TAT) polypeptide PRO62893, SEQ:652.
XX KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic.
XX OS Homo sapiens.
XX PN WO2004030615-A2.
XX PD 15-APR-2004.
XX PF 29-SEP-2003; 2003WO-US028547.
XX PR 02-OCT-2002; 2002US-0414971P.
XX PA (GETH) GENENTECH INC.
XX PI Wu TD, Zhang Z, Zhou Y;
XX DR WPI: 2004-347921/32.
XX DR N-PSDB; ACN37637.
XX PT New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.

XX

PS Claim 12; SEQ ID NO 652; 7273pp; English.

CC The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT polypeptide of the invention

SQ Sequence 389 AA;

Alignment Scores:

Pred. No.:	1.42e-165	Length:	389
Score:	1454.00	Matches:	269
Percent Similarity:	94.08%	Conservative:	33
Best Local Similarity:	83.80%	Mismatches:	19
Query Match:	78.76%	Indels:	0
DB:	8	Gaps:	0

US-10-698-228-12 (1-1053) x ABM80258 (1-389)

QY 91 GAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTCCATCCAGTACCCT 150

Db 69 AspGluProLeuLeuArgGluAsnProArgargPheValIlePheProIleGluTrpHis 88

QY 151 GATATTTGGAAATGTATAAACAGGCACAGCTTCTCTCGACAGCAGAGAGGTTGAC 210

Db 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValasp 108

QY 211 TTATCAAAGATCTCCCTACTGTGAACAAGCTTAAAGCAGATGAGAACTTCTCTCT 270

Db 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128

QY 271 CACATCTTAGCCTTTTTCAGCCAGATGATGGAATTTGAAATGAAATTTGGTGGAGCGC 330

Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148

QY 331 TTTAGTCAGGAGGTGTCAGGTTCCAGAGCTCGCTGTCTTCTATGGCTTTCAAAATTCATC 390

Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168

QY 391 GAGAATGTTCACTCAGAGATGTACAGTTTGTGTATAGACACTTACATCAGAGATCCCAAG 450

Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188

QY 451 AAAAGGGAATTTTATTTAATGCAATTTGAAACCACTGCCCTATGTTAAAGAAAACAGAT 510

Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208

QY 511 TGGGCTTGGATGGATAGCAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGCCTTT 570

Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValValAlaPhe 228

QY 571 GCTGCTGTAGAAGAGTTTCTTCTCAGGATCTTTTCTGCTATATTTCTGCTATAATCTGCT 630

Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248

QY 631 AGAGTCTTATGCGCAGGACTCACTTTTCCNATGAACATCATCAGCAGAGATGAAGACTT 690


```

Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 368
QY 691 CACTGTGACTTTGCTGCTGATGTTCCAACTACTAGTAATAAGCCTTCAGAGAAGG 750
Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGAGATCATTTGTCATGCTGTCACAAATTTGACGAGAGTTTAAACAGAGCCTTG 810
Db 289 ValArgGluIleIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCCTCATTTGAAGTAATGCAATTTTGATGAACAGPACATTTGAGTTGTAGCT 870
Db 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnIleGluPheValAla 328
QY 871 GACAGATTAATGTCGAACTGCTCAAGTCTCAAGTCTTTCAGCAGAAATCCTTTGAT 930
Db 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY 931 TTTATGGAACCAATTTCTTTAGAGGAAACCAAAATTTCTTTCAGAAACGAGTTTCAGAG 990
Db 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY 991 TATCAGCGTTTTCAGTTATGCGAGAAACCAACGATATACGCTCTTACCTTGGATGCAGAT 1050
Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY 1051 TTT 1053
Db 389 Phe 389
RESULT 14
AAB43902
ID AAB43902 standard; protein; 413 AA.
XX
AC AAB43902;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human cancer associated protein sequence SEQ ID NO:1347.
XX
KW Human; cancer associated gene; cancer antigen; detection; cancer;
KW diagnosis; cytostatic; proliferative; vulnerable; immunomodulator;
KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW antinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW dermatological; neuroprotective; thrombolytic; coagulant; neoplasia;
KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW allergic reaction; graft versus host disease; organ rejection;
KW haemostatic; thrombolytic; cardiovascular disorder; infection;
KW neurological disease; drug screening.
XX
OS Homo sapiens.
XX
PN WO200055350-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US005882.
XX
PR 12-MAR-1999; 99US-0124270P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI: 2000-587533/55.
XX
DR N-PSDB; AAC78111.
XX
PT Novel isolated nucleic acids comprising sequences encoding peptides
PT useful for treating or diagnosing e.g. cancer.
XX
XX Claim 11; Page 2000-2002; 2352pp; English.
PS

```

```

XX AAC77607 to AAC78448 encode the human cancer associated proteins given in
CC AAB43998 to AAB44239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vulnerable; immunomodulator;
CC antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC antinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC vasotropic; antipsoriatic and antiangiogenic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention
XX
SQ Sequence 413 AA;
Alignment Scores:
Pred. No.: 1-46e-165 Length: 413
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 3 Gaps: 0

```

```

US-10-698-228-12 (1-1053) x AAB43902 (1-413)
QY 91 GAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTTCTCATCTTTTCCAAATCCAGTACCCT 150
Db 93 AspGluProLeuLeuArgGluAsnProArgArgPheValIlePheProIleGluTyrHis 112
QY 151 GATATTGGAAATGTATAAACAGGACAGGCTTCCTCTGGACAGCAGAGAGGTTGAC 210
Db 113 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 132
QY 211 TTATCAAGAGTCTCCCTCACTCGAACAGCTTAACAGCATGAGAGTACTTCATCTCT 270
Db 133 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 152
QY 271 CACATCTTAGCCTTTTTCAGCCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 330
Db 153 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 172
QY 331 TTTAGTCAGAGGTGCAGGTTCAGAGGTTCAGAGGTTCAGAGGTTCAGAGGTTCAGAGGTTC 390
Db 173 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 192
QY 391 GAGAAATGTTTACCTCAGAGATGACAGTTTCTGATACAGACTTACATCAGAGTCCCAAG 450
Db 193 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 212
QY 451 AAAAGGGAATTTTATTATTAATGCAATTAATGCAATTAATGCAATTAATGCAATTAATGCA 510
Db 213 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 232
QY 511 TGGGCGCTTGCATGAGATGAGATGAGAAATCTACTTTTGGGGAAGAGTGGTGGCCTTT 570
Db 233 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaAlaPhe 252
QY 571 GCTGCTGTAGAAGAGTTTCTTCTCAGAGTCTTTTCTGCTGCTATATTCTGCTAAAGAG 630
Db 253 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 272
QY 631 AGAGGTCTTATGCCAGGACTCACTTTTCCAACTCACTATCAGCAGAGATGAAGACTTT 690

```

Db 273 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuLeuSerArgAspGluGlyLeu 292
 QY 691 CACTGTGACTTTGCTTGGCTGTATGTTCCCAATCTAGTAATAAGCCTTCAGAGAAGG 750
 Db 293 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 312
 QY 751 GTCAGGAGATCATTTGCTGATGCTGCTCAAAATTTGAGCAGGAGTTTAAACAGAGCCTTG 810
 Db 313 ValArgGluIleIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 332
 QY 811 CCAGTTGGCTCATTCGATGATTCATTTGATGAACAGTACATTTGAGTTTGTAGCT 870
 Db 333 ProValLeuLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 352
 QY 871 GACAGATTACTTGTGAATCTGATCTCAAGGTTTTCAGGCAGAAATCTTTTGAT 930
 Db 353 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 372
 QY 931 TTTATGAAAAATTTCTTTAGAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAGAG 990
 Db 373 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 392
 QY 991 TATCAGGCTTTGTCAGTTATGTCAGAAACACAGATAAGCTCTTACCTTGGATGAGAT 1050
 Db 393 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 412
 QY 1051 TTT 1053
 Db 413 Phe 413
 RESULT 15
 AAU28017
 ID AAU28017 standard; protein; 453 AA.
 AC AAU28017;
 XX
 DT 18-DEC-2001 (first entry)
 DE Human contig polypeptide sequence #170.
 XX
 KW Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
 KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
 KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
 KW nervous system disorder; inflammatory disorder; cell differentiation;
 KW angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
 KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;
 KW cytostatic; antirheumatic; antiarthritic; vulnery; antiinflammatory;
 KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
 KW neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;
 KW immunostimulant; analgesic; gene therapy.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO200164834-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 26-FEB-2001; 2001WO-US004926.
 XX
 PR 28-FEB-2000; 2000US-00515126.
 PR 18-MAY-2000; 2000US-00577409.
 PR 17-JUN-2000; 2000US-00597707.
 PR 14-JUL-2000; 2000US-00616807.
 PR 19-SEP-2000; 2000US-00664641.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;
 PI Drmanac R;
 XX
 XX WPI; 2001-589862/66.

DR N-PSDB; AAS44917.
 XX Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis, treatment of cancer,
 PT neurological, inflammatory disorders and for use in arrays for detection.
 XX
 PS Claim 10; Page 146-147; 153pp; English.

XX Sequences AAU27676-AAU28019 represent full-length polypeptides and contig
 CC polypeptides of the invention. The proteins and their associated DNA
 CC sequences are useful for the treatment, diagnosis and prevention of
 CC various types of disorder in a mammalian subject such as a human, dog,
 CC monkey, mouse, hamster or rat. The disorders include cancers such as
 CC leukemia, lymphoma and neuroblastoma, autoimmune disorders such as
 CC multiple sclerosis, connective tissue disease, rheumatoid arthritis,
 CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
 CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
 CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
 CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's
 CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
 CC bowel disease. The sequences exhibit activity relating to angiogenesis,
 CC cell proliferation, cell differentiation, stem cell growth factor,
 CC activin or inhibin. Therefore, they can be used to manipulate stem cells
 CC in culture to give rise to neuroepithelial cells that can be used to
 CC augment or replace cells damaged by illness, accidental damage or genetic
 CC disorders. The sequences may also be used for regeneration of bone,
 CC cartilage, tendons and ligaments and in tissue repair and burn healing.
 CC Note: Some sequences for this patent did not form part of the printed
 CC specification, but were obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 453 AA;

Alignment Scores:

Pred. No.: 1.53e-165 Length: 453
 Score: 1454.00 Matches: 269
 Percent Similarity: 94.08% Conservative: 33
 Best Local Similarity: 83.80% Mismatches: 19
 Query Match: 78.76% Indels: 0
 DB: Gaps: 0

US-10-698-228-12 (1-1053) x AAU28017 (1-453)

QY 91 GAAGAGCCACTCTTAAGAAAGAGTCTCGCGGTTTGTCTATCTTCCATTCAGTACCT 150
 Db 133 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 152
 QY 151 GATATTGGAAAATGTATAACAGGCACAGCTTCTCTGGACAGCAGAGAGGTTGAC 210
 Db 153 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 172
 QY 211 TTATCAAGGATCTCCCTCACTGCAACAAGCTTAAAGCAGATGAGAGTACTTCTCTCT 270
 Db 173 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 192
 QY 271 CACATCTTAGCCTTTTTCAGCCAGTGTGAATTAATAATGAAATTTGGTGAGCGC 330
 Db 193 HisValLeuAlaPhePheAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 212
 QY 331 TTTAGTCAGGAGGTTCAGGTTCCAGAGCTCGCTTCTATGCTTCTATGCTTCAATTTCTCATC 390
 Db 213 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 232
 QY 391 GAGAATGTTTCACTCAGAGATGTACAGTTTGTCTATAGACACTTACATCAGAGATCCCAG 450
 Db 233 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 252
 QY 451 AAAAGGAAATTTTATTTAATGCAATTAATGAAACCATGCCCTTATGTTAAGAAAAAGCAGAT 510
 Db 253 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAsp 272
 QY 511 TGGGCTTCGGATGAGATAGATAGAAATCTACTTTTGGGAAAGAGTGTGCGCTTT 570

Db 273 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValValAlaPhe 292
QY 571 GCTGCTGTAGAGGAGTTTCTTCTCAGGATCTTTTGGCTGTATATTTCTGGCTAAAGAG 630
Db 293 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 312
QY 631 AGAGGTCTTATGCCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGGACTT 690
Db 313 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 332
QY 691 CACTGTGACTTTGCTTGCCTGATGTTCCAAATAGTAAATAGCCTTCAGAAAGG 750
Db 333 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 352
QY 751 GTCAGGAGATCATTTGCTGTCGTCAAAATTGACGAGGAGTTTTCACAGAGCCTTG 810
Db 353 ValArgGluIleIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 372
QY 811 CCAGTTGGCCTCATTTGGAATGAATTGCATTTTCATGAACAGTACATTCAGTTTGTAGCT 870
Db 373 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 392
QY 871 GACAGATTACTTGTGGAACCTTGATTCCTCAAGGTTTTCAGGCAGAAAATCCTTTTGAT 930
Db 393 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 412
QY 931 TTTATGGAAAACATTTCTTAGAGGAAAAACAATTTCTTTCAGAAACGAGTTTCAGAG 990
Db 413 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 432
QY 991 TATCAGCGTTTTCGAGTTATGGCAGAAACACAGATAAGTCTTCACCTTGGATGCAGAT 1050
Db 433 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 452
QY 1051 TTT 1053
Db 453 Phe 453

Search completed: October 30, 2005, 06:28:52
Job time : 73.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 30, 2005, 05:15:40 ; Search time 15.75 Seconds
(without alignments)
12865.556 Million cell updates/sec

Title: US-10-698-228-12

Perfect score: 1846

Sequence: 1 atggcgaccggaaggcc.....tcacttgatgcagatttt 1053

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ n2p.model -DEV=xlh
-O=/cgn2_1/USPTO_spo1/US10698228/runat_26102005_100608_3995/app_query.fasta_1.2446
-DB=PIR_79 -QFMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=500 -THR_MIN=0 -THR_MAX=100 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228@cgn_1_1_39 @runat_26102005_100608_3995 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THRADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	98.6	366	T46249	hypothetical prote
2	1454	78.8	389	S25854	ribonucleoside-dip
3	1443	78.2	390	S26735	ribonucleoside-dip
4	1421	77.0	386	S27153	ribonucleoside-dip
5	1331	72.1	384	S24585	ribonucleoside-dip
6	1280	69.3	319	T30782	ribonucleoside-dip
7	1280	69.3	319	B27154	E4L protein - vari
8	1279	69.3	319	RDVZV	ribonucleoside-dip
9	1279	69.3	333	T28466	ribonucleoside-dip
10	1266	68.6	333	H36839	C8L protein - vari
11	1258.5	68.2	348	T30470	ribonucleoside-dip
12	1217	65.9	299	RDSS2R	ribonucleoside-dip
13	1214	65.8	381	T18876	hypothetical prote
14	1169.5	63.4	399	A26916	ribonucleoside-dip

15	1161.5	62.9	391	2	T39992	ribonucleoside-dip
16	1161.5	62.9	391	2	S34808	ribonucleoside-dip
17	1128.5	61.1	329	2	T03688	ribonucleoside-dip
18	1070.5	58.0	340	2	S68538	ribonucleoside-dip
19	1024.5	55.5	349	2	B49412	ribonucleoside-dip
20	950	51.5	324	2	T17978	probable ribonucle
21	805	43.6	345	2	S59744	ribonucleoside-dip
22	681	36.9	668	2	T29884	hypothetical prote
23	638	34.6	327	1	RDVZAS	ribonucleoside-dip
24	385	20.9	331	2	A84389	ribonucleoside red
25	384.5	20.8	415	2	A83502	ribonucleoside red
26	358.5	19.4	346	2	G81728	ribonucleoside-dip
27	357.5	19.4	346	2	E71466	probable ribonucle
28	335	18.1	346	2	B72010	ribonucleoside-dip
29	335	18.1	346	2	F86613	ribonucleoside red
30	329	17.8	306	1	WMBE18	ribonucleoside-dip
31	324	17.6	303	2	T47526	ribonucleoside-dip
32	322	17.4	320	2	T42563	ribonucleoside-dip
33	318	17.2	321	1	WMBE41	ribonucleoside-dip
34	302.5	16.4	314	1	WMBE84	ribonucleoside-dip
35	301	16.3	302	1	WMBE12	ribonucleoside-dip
36	294.5	16.0	347	2	C87281	hypothetical prote
37	292.5	15.8	340	1	WMBE57	ribonucleoside-dip
38	290	15.7	305	2	S55655	ribonucleoside-dip
39	290	15.7	328	2	C97781	hypothetical prote
40	289.5	15.7	337	1	WMBE32	ribonucleoside-dip
41	288.5	15.6	337	1	WMBE82	ribonucleoside-dip
42	288.5	15.6	341	2	C71858	ribonucleoside-dip
43	286.5	15.5	341	2	D64565	ribonucleoside-dip
44	285	15.4	305	2	T03155	ribonucleoside-dip
45	285	15.4	324	2	B71655	ribonucleoside-dip

ALIGNMENTS

RESULT 1

T46249
hypothetical protein DKFZp761E1312.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C;Accession: T46249
R;Ansoorge, W.; Winkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23028
A;Accession: T46249
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-366 <AAA>
A;Cross-references: UNIPROT:Q9NTD8; EMBL:AL137348
A;Experimental source: adult amygdala; clone DKFZp761E1312
C;Genetics:
A;Note: DKFZp761E1312.1
C;Superfamily: ribonucleoside reductase small subunit

Alignment Scores:
Pred. No.: 1.57e-157 Length: 366
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 2 Gaps: 0

US-10-698-228-12 (1-1053) x T46249 (1-366)

QY 1 ATGGGGAGCCCGAAAGCGCGCGCTGGATCAGGATGAGATCATCTTCA 60

Db 16 MeGlyaspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 35

QY 61 GACACCAACGAAGTGAATAAAGTCAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120

Db 36 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 55

QY 121 CGGTTTGTCATCTTTCCATCCAGTACCTCGATATTGGAAAATGTATAAACAGGCACAG 180

Db 56 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 75
QY 181 GCTTCTCTCTGACACAGAGGGTGTACTTATCAAGGATCTCCCTCACTGGACAG 240
Db 76 AlaSerPheTrpTrpAlaGluGluValAlaSerLysLeuSerLysAspLeuProHisTrpAsnLys 95
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTACATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db 96 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 115
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCCCTTTAGTCAGAGGTGCAGTTCCAGAGGCT 360
Db 116 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 135
QY 361 CGCTGTTCTATGGCTTTCAAATCTCATCGAGAATGTTCACTCAGAGATGACAGTTTG 420
Db 136 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 155
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAATTTTATTTAATGCAATTGAA 480
Db 156 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 175
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATGGCCCTTCGGATGGATGATAGCAAGAAA 540
Db 176 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 195
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCCCTTCTGCTGTAGAGGAGTTTCTCTCAGGA 600
Db 196 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 215
QY 601 TCTTTTGTCTATATTCTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
Db 216 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 235
QY 661 AATGAACCTCATCAGCAGATGAAGGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 720
Db 236 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 255
QY 721 TACTTAGTAATAAGCCTTCAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 256 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 275
QY 781 ATTGACGAGGTTTTTACAGAAGCCTTGCAGTGGCTCATTTGGAATGATTCATT 840
Db 276 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 295
QY 841 TTGATGAAACAGTACATTGAGTTGTAGCTGCAGAGATTACTTGTGAACTTGGATTCTCA 900
Db 296 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 315
QY 901 AAGTTTTTTCAGCAGAAATCCTTTTGAATTTATGGAACAACTTTTTCAGAGAA 960
Db 316 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 335
QY 961 ACAATTTCTTTCAGAAACAGTTTCAGAGTATCAGCTTTTTCAGTTATGGCAGAAACC 1020
Db 336 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 355
QY 1021 ACAGATAACGCTTCTACCTTGGATGCAGATTTT 1053
Db 356 ThrAspAsnValPheThrLeuAspAlaAspPhe 366

RESULT 2
S25854
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - human
N:Alternate names: ribonucleotide reductase M2 chain; ribonucleotide reductase small chain
C:Species: Homo sapiens (man)
C:Date: 08-Jun-1994 #8sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
R:Pavloff, N.; Rivard, D.; Masson, S.; Shen, S.H.; Mes-Masson, A.M.
C:Accession: S25854
DNA Seq. 2, 227-234, 1992
A:Title: Sequence analysis of the large and small subunits of human ribonucleotide reductase

A:Reference number: S25853; MUID:92329977; PMID:1627826
A:Accession: S25854
A:Molecule type: mRNA
A:Residues: 1-389 <P>
A:Cross-references: UNIPROT:P31350; EMBL:X59618; NID:G36154; PIDN:CAA42181.1; PID:G36155
C:Genetics:
A:Gene: GDB:RRM2
A:Cross-references: GDB:120358; OMIM:180390
A:Map position: 2p25-2p24
C:Superfamily: ribonucleoside reductase small subunit
C:Keywords: deoxyribonucleotide biosynthesis; iron; metalloprotein; oxidoreductase
F:138,169,172,233,266,269,269Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, His) #str
F:176/Active site: Tyr (stable tyrosyl radical) #status predicted
Alignment Scores:
Pred. No.: 4,1e-124 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 2 Gaps: 0
US-10-698-228-12 (1-1053) x S25854 (1-389)
QY 91 GAAAGAGCCACTCTCTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTTCCAAATCCAGTACCCT 150
Db 69 AspGluProLeuLeuArgGluAsnProArgArgPheValIlePheProIleGluTyrHis 88
QY 151 GATATTGGAAATGATATAAACAGGACAGGCTTCTCTGGACAGCAGAGAGAGGTTGAC 210
Db 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpTrpAlaGluGluValAsp 108
QY 211 TTATCAAGGATCTCCCTCACTCGAACACAGCTTAAAGCAGATGAGAGTACTTCTCATCT 270
Db 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTTAGCCCTTTTTCAGCCAGGATGGAATTTGAAATGAAATTTTCGTGGAGCGC 330
Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY 331 TTTAGTCAGAGGTGCAGGTTCCAGAGCTCGCTGTTTCTTCTAGCTTTTCAAAATCTCATC 390
Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGAAATGTTCACTCAGAGATGTACAGTTTCTGTATAGACACTTACATCAGAGATCCCAAG 450
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAAGGGAATTTTATTAAATGCAATGAAACCATGCCCTATGTTAAAGAAAAAGCAGAT 510
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAsp 208
QY 511 TGGGCCCTTCGATGGATGATAGAGATCTACTTTTGGGAAAGAGTGGTGGCCCTTT 570
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaAlaPhe 228
QY 571 GCTGCTGTAGAAGAGTTTCTCTCAGATCTTTTCTGCTATATTTCTGCTGCTAAAGAG 630
Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGGTCTTATGCCAGGACTCACTTTTCCAACTCACTCAGCAGATGAAGACTT 690
Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY 691 CACTGTGACTTTTCTGCTGATGTTCCAACTTCTAGTAATAAGCCCTTCAGAGAAAGG 750
Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGAGATCATTTGTTGATGCTGTCAAAATTCAGAGAGGATTTTAAACAGAGCCTTG 810
Db 289 ValArgGluIleIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCCCTCATTTGGAATGAAATTTGATTTGATGAACAGTACATTTGAGTTGTAGCT 870

```

Db      309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
QY      871 GACAGATTACTTGTGGAACCTTGATTTCTCAAGGTTTTCAGGCAGAAATCCTTTTGAT 930
Db      329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY      931 TTTATGGAACAACTTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAGAG 990
Db      349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY      991 TATCAGCGTTTTCAGTTATGCGAGAAACACAGATAACGCTTTCACCTTGGATGCGAT 1050
Db      369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY      1051 TTT 1053
Db      389 Phe 389

RESULT 3
S06735
ribonucleoside-diphosphate reductase (EC 1.17.4.1) chain M2 - mouse
N:Alternate names: ribonucleotide reductase M2 subunit
C:Species: Mus musculus (house mouse)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S06735; A24835
R:Thelander, M.; Thelander, L.
EMBO J. 8, 2475-2479, 1989
A:Title: Molecular cloning and expression of the functional gene encoding the M2 subunit
A:Reference number: S06735; MUID:90060004; PMID:2684652
A:Accession: S06735
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-390 <TH>
A:Cross-references: UNIPROT:P11157; EMBL:X15666; NID:g50719; PIDN:CAA33707.1; PID:g50720
R:Thelander, L.; Berg, P.
Mol. Cell. Biol. 6, 3433-3442, 1986
A:Title: Isolation and characterization of expressible cDNA clones encoding the M1 and M2 subunits of mouse ribonucleoside diphosphate reductase
A:Reference number: A24835; MUID:87089677; PMID:3025593
A:Accession: A24835
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-390 <TH>
A:Cross-references: GB:M14223; NID:g200767; PIDN:AAA40062.1; PID:g200768
A>Note: the authors translated the codon GGG for residue 315 as Glu
C:Genetics:
A:Map position: 12
A:Introns: 33/3; 58/3; 107/3; 146/3; 191/2; 223/1; 267/3; 302/3; 340/3
C:Superfamily: ribonucleoside reductase small subunit
C:Keywords: iron; oxidoreductase
F:139,170,173,233,267,270/Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, His) #st
F:177/Active site: Tyr (stable tyrosyl radical) #status predicted

Alignment Scores:
Pred. No.: 4, 12e-123 Length: 390
Score: 1443.00 Matches: 269
Percent Similarity: 90.62% Conservative: 40
Best Local Similarity: 78.89% Mismatches: 26
Query Match: 78.17% Indels: 6
DB: 1 Gaps: 1

US-10-698-228-12 (1-1053) x S06735 (1-390)

QY      49 AGATCATCTTCAGACACCAAGAGTGAATAAAGTCA-----AAT 90
Db      50 ArgIlePheGlnAspSerAlaGluLeuGluSerLysAlaProThrAsnProSerValGlu 69
QY      91 GNAGAGCACTCTTGAAGAAGATTCTCGCCGGTTTGTCATTTTCCAATCCAGTACCTT 150
Db      70 AspGluProLeuArgGluAsnProArgPheValPheProIleGluTyrHis 89
QY      151 GATATTTCGAAATCTATAAACAGGCACACAGGCTTCTCTTCGACAGACAGAGTTGAC 210

```

```

Db      90 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheThrAlaGluGluValAsp 109
QY      211 TTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTCTCTCT 270
Db      110 LeuSerLysAspIleGlnHisTrpGluAlaLeuLysProAspGluArgHisPheIleSer 129
QY      271 CATCATCTTAGCCTTTTTCAGCCAGCTGATGGAATTTGAAATGAAAAATTTTGGTGAGCCG 330
Db      130 HisValLeuAlaPhePheAlaAlaSerAspGlyLeValAsnGluValGluArg 149
QY      331 TTTAGTCAGAGGTGCGAGTTTCAGAGCTCGCTGTTTCTATGGCTTTCAAAATTCATC 390
Db      150 PheSerGlnGluValGlnValThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 169
QY      391 GAGNATCTTCACTCAGAGATGTACAGTTTGTGTAGACACTTACATCAGAGATCCCAG 450
Db      170 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 189
QY      451 AAAAGGGAATTTTATTTAATGCAATGAAACCATGCCCTATGTTAAAGAAAAAGCAGAT 510
Db      190 GluArgGlnTyrLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 209
QY      511 TGGCCCTTGCATGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGGTGCCTTT 570
Db      210 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 229
QY      571 GCTGCTGTAGAAGAGTTTCTTCTCAGGATCTTTTGTCTGTATATTCTGGCTAAAGAAG 630
Db      230 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLys 249
QY      631 AGAGGTCTTATGCGAGGACTCACCTTTTCCAACTCACTCAGCAGAGATGAAGACAT 690
Db      250 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluLeu 269
QY      691 CACTGTGACTTTGCTTGCCTGATGTTCCAACTTAGTAAATAGCCCTTCAGAGAAAGG 750
Db      270 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProAlaGluGlnArg 289
QY      751 GTCAGGAGATCATTTGTTGATGCTGTCAAAATGAGCAGAGAGTGTTTTAAACAGAGCCTTG 810
Db      290 ValArgGluIleIleThrAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 309
QY      811 CCAGTTGGCCTCATTTGGAATGAAATGCAATTTGATGAAACAGTACATTTAGTTTGTAGCT 870
Db      310 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 329
QY      871 GACAGATTACTTGTGGAACCTTGATTTCTCAAGGTTTTCAGGCAGAAATCCTTTTGAT 930
Db      330 AspArgLeuMetLeuGluLeuGlyPheAsnLysIlePheArgValGluAsnProPheAsp 349
QY      931 TTTATGGAACAACTTTCTTTAGAGGAAACAAATTTCTTTTGAAGAACGAGTTTCAGAG 990
Db      350 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 369
QY      991 TATCAGCGTTTTCAGTTATGCGAGAAACACAGATAACGCTTTCACCTTGGATGCGAT 1050
Db      370 TyrGlnArgMetGlyValMetSerAsnSerThrGluAsnSerPheThrLeuAspAlaAsp 389
QY      1051 TTT 1053
Db      390 Phe 390

```

RESULT 4

S27153
 ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - golden hamster
 N:Alternate names: ribonucleotide reductase small chain
 C:Species: Mesocricetus auratus (golden hamster)
 C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
 C:Accession: S27153
 R:Chaudhuri, M.M.; Tonin, P.N.; Srinivasan, P.R.
 Biochim. Biophys. Acta 1171, 117-121, 1992
 A:Title: cDNA sequence of the small subunit of the hamster ribonucleotide reductase.
 A:Reference number: S27153; MUID:93042000; PMID:1384717

A;Accession: S27153
A;Molecule type: mRNA
A;Residues: 1-386 <CHA>
A;Cross-references: UNIPROT:Q60561; EMBL:X68127; NID:g49649; PIDN:CAA48232.1; PID:g49650
C;Superfamily: ribonucleoside reductase small subunit
C;Keywords: oxidoreductase

Alignment Scores:
Pred. No.: 4,148-121 Length: 386
Score: 1421.00 Matches: 286
Percent Similarity: 91.57% Conservative: 38
Best Local Similarity: 80.12% Mismatches: 20
Query Match: 76.98% Indels: 8
DB: 2 Gaps: 2

US-10-698-228-12 (1-1053) x S27153 (1-386)

QY 70 GAAGTGAATAAAGTCAAAT-----GAAGAGCCACTCTCTAAGAAAGAGTCTT 117
Db 59 GluSerLysValSerThrAsnProSerValGluAspGluProLeuLeuArgGluAsnPro 78
QY 118 CCGCGTTTGTTCATCTTCCATCCAGTACCCTGATATTGGAAAATGTATAAACAGGCA 177
Db 79 ArgArgPheValValPheProIleGluTyrHisAspIleTrpLysMetTyrLysLysAla 98
QY 178 CAGGCTTCCTCTGACACAGAGGTTGACTTATCAAGGATCTCCCTCAGTGGAAC 237
Db 99 GluAlaSerPheTrpThrAlaGluValAspLeuSerLysAspIleGlnHisTrpGlu 118
QY 238 AAGCTTAAAGCAGATGAGAAGTACTCTCATCTCTCAGATCTAGCCCTTTTTCAGCCAGT 297
Db 119 AlaLeuLysProAspGluArgHisPheIleSerHisValLeuAlaPhePheAlaAlaSer 138
QY 298 GATGGAATTGTAATAAATTTGGTGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAG 357
Db 139 AspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValThrGlu 158
QY 358 GCTCGCTGTTCTATGGCTTCCAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGT 417
Db 159 AlaArgCysPheTyrGlyPheGlnIleAlaMetGluAsnIleHisSerGluMetTyrSer 178
QY 418 TTGCTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATT 477
Db 179 LeuLeuIleAspThrTyrIleLysAspSerLysGluArgGluTyrLeuPheAsnAlaIle 198
QY 478 GAAACATGCCCTATGTTTAAGAAAAAGCAGATTGGCGCTTCGGATGGATAGCAGATAG 537
Db 199 GluThrMetProCysValLysLysLysAlaAspTrpAlaLeuArgTrpIleGlyAspLys 218
QY 538 AAATCTACTTTTGGGGAAGAGTGGTGGCTTTCGCTGTAGAGGAGTTTCTCTCA 597
Db 219 GluAlaThrTyrGlyGluArgValValAlaPheAlaAlaValGluGlyIlePhePheSer 238
QY 598 GGATCTTTTGTGCTATATTCTGGCTTAAAGAGAGAGGTCTTATGCGAGGACTCACTTTT 657
Db 239 GlySerPheAlaSerIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPhe 258
QY 658 TCCAAATGAATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTCCCTCGATGTT 717
Db 259 SerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPhe 278
QY 718 CAATACTAGTAAATGAAGCTTCAGAAAGAGGTGAGGAGATCATTTGTTGATGCTGTC 777
Db 279 LysHisLeuValHisLysProSerGluGlnArgValGlnGluIleIleThrAsnAlaVal 298
QY 778 AAAATTGACGAGGAGTTTTTACAGAGAGCTGCCAGTTGGCTCATTCGGAATGAATTGC 837
Db 299 ArgIleGluGlnGluPheLeuThrGluAlaLeuProValLysLeuIleGlyMetAsnCys 318
QY 838 ATTTTTCATGAACAGTACATTCAGTTTGTAGCTGCAGAGATTAATCTGTGGAACCTTGGATTC 897
Db 319 ThrLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuMetLeuGluLeuGlyPhe 338

QY 898 TCAAAGCTTTTTCAGGCAGAAAAATCCTTTTATGATTTTATGAAAAACATTTCTTTAAGGA 957
Db 339 AsnLysIlePheLysValGluAsnProPheAspPheMetGluAsnIleSerLeuGluGly 358
QY 958 AAAACAAAATTTCTTTGAGAAACAGATTTCAGAGTATCAGCGTTTTCGACGTTATGCGCAGAA 1017
Db 359 LysThrAsnPhePheGluLysArgValGlyGluTyrGlnArgMetGlyValMetSer--- 377
QY 1018 ACCACAGATAACGCTCTTCACCTTGGATGACAGATTTT 1053
Db 378 -----AsnSerPheThrLeuAspAlaAspPhe 386

RESULT 5
S24585
ribonucleoside-diphosphate reductase (EC 1.17.4.1) chain M2 - Atlantic surf clam
N;Alternate names: ibonucleoside-diphosphate reductase small chain
C;Species: Spisula solidissima (Atlantic surf clam)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S24585
R;Hunt, T.; Standart, N.
submitted to the EMBL Data Library, November 1990
A;Reference number: S24584
A;Accession: S24585
A;Molecule type: mRNA
A;Residues: 1-384 <HUN>
A;Cross-references: UNIPROT:P07201; EMBL:X55125; NID:g10341; PIDN:CAA38919.1; PID:g10342
C;Superfamily: ribonucleoside reductase small subunit
C;Keywords: oxidoreductase

Alignment Scores:
Pred. No.: 6,458-113 Length: 384
Score: 1331.00 Matches: 257
Percent Similarity: 85.26% Conservative: 38
Best Local Similarity: 74.28% Mismatches: 41
Query Match: 72.10% Indels: 10
DB: 2 Gaps: 4

US-10-698-228-12 (1-1053) x S24585 (1-384)

QY 46 GAGAGATCATCTTCAGACACCAAC---GAAAGTGAATAAAG----- 84
Db 39 GlnArgSerThrGlnLysThrProLeuLysGlnGluIleLysProValValLysLysSer 58
QY 85 TCAAAATGAAGAGCCACTCTTAAGAAAGAGTCTCGCGGTTTGTTCATCTTTCCAAATCCAG 144
Db 59 GlnGlnValGluProLeuLeuAlaAspAsnProArgArgPheValValLeuProIleGln 78
QY 145 TACCTCTGATATTTGGAAAAATGTATAACAGCAGCAGGCTTCCTTCTGACAGCAGAGAAG 204
Db 79 TyrHisAspIleTrpLysMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGlu 98
QY 205 GTTGACTTATCAAAGGATCTCCCTCACTGGAAACAGCTTAAAGCAGATGAGAAAGTACTTC 264
Db 99 ValAspLeuSerLysAspMetAlaHisTrpGluSerLeuLysLysGluGluLysHisPhe 118
QY 265 ATCTCTCAGATCTTAGCCTTTTTCGAGCCAGTGTGAATTTGTAATGAAATTCGTTG 324
Db 119 IleSerHisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuVal 138
QY 325 GAGCGCTTTAGTCAGGAGGTGCGAGGTTCAGAGGCTCGCTGTGTTTCTATGGCTTTCAAAT 384
Db 139 GluArgPheSerLysGluValGlnValThrGluAlaArgCysPheTyrGlyPheGlnIle 158
QY 385 CTCATCAGAAATGTCTCATCAGAGATGTACAGTTTGTGTGATAGACACTTACATCAGAGAT 444
Db 159 AlaMetGluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAsp 178
QY 445 CCCAAGAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCGCTTATGTTAAGAAAAAA 504
Db 179 ProGlnGluArgAspPheLeuPheAsnAlaIleGluThrMetProCysValLysGluLys 198
QY 505 GCAGATTGGCGCTTGGATGATAGCAGATAGAAAAATCTACTTTTGGGGAAGAGTGTG 564

Db 199 AlaAspTrpAlaMetArgTrpIleAsnAspAspSerSerSerTyrAlaGluArgValVal 218
QY 565 GCCTTGTCTGTAGAACGAGTTTCTCTCAGGATCTTTGCTGCTATATCTCGCTA 624
Db 219 AlaPheAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeu 238
QY 625 AGAAGAGAGGCTTTATGTCAGGACTCACTTTTCCAACTCACTCAGCAGAGATGAA 684
Db 239 LysLysArgGlyIleMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGlu 258
QY 685 GGACTTCATCTGACTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 744
Db 259 GlyLeuHisCysAspPheAlaCysLeuMetPheSerHisLeuValAsnLysProSerGln 278
QY 745 GAAAGGTCAGGAGATCAATGTTGATGCTGCTCAAAATTTGAGCAGGAGTTTAAACAGAA 804
Db 279 GluArgIleHisGlnIleIleAspGluAlaValLysIleGlnValPheLeuThrGlu 298
QY 805 GCCTTGCCAGTTGGCTCAATGGAATGGAATGCAATTTGATGAAACAGTACATTTAGTTT 864
Db 299 AlaLeuProCysArgLeuIleGlyMetAsnCysAspLeuMetArgGlnTyrIleGluPhe 318
QY 865 GTAGCTGACAGATTAATCTGGAACCTGGAATCTCAAGGTTTTCAGCAGCAAAATCTT 924
Db 319 ValAlaAspArgLeuLeuLeuLysCysAspLysLeuTyrAsnLysGluAsnPro 338
QY 925 TTTGATTTTATGAAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACAGATT 984
Db 339 PheAspPheMetGluHisIleSerLeuGluGlyLysThrAsnPhePheGluLysArgVal 358
QY 985 TCAGAGTATCAGGTTTTCAGTTATGCA-----GAAACACACAGAT---AACGCTCTC 1035
Db 359 GlyGluTyrGlnLysMetGlyValMetSerGlyGlyAsnThrGlyAspSerHisAlaPhe 378
QY 1036 ACCTTGATGTCAGATTTT 1053
Db 379 ThrLeuAspAlaAspPhe 384
RESULT 6
T30782
Ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - vaccinia virus (strain
N1) Alternate names: ribonucleotide reductase, small subunit
C:Species: vaccinia virus
A:Variety: strain Ankara
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
R:Antoine, G.; Scheiflinger, F.; Falkner, F.G.; Dorner, P.
Submitted to the EMBL Data Library, March 1997
A:Description: The complete genomic sequence of the Modified Vaccinia Ankara (MVA) strain
A:Reference number: Z20877
A:Accession: T30782
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-319 <ANT>
A:Cross-references: UNIPROT:O57175; EMBL:U94848; PIDN:AB96415.1
A:Experimental source: strain Ankara
C:Genetics:
A:Note: MVA032L
C:Function:
A:Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleoside
C:Superfamily: ribonucleoside reductase small subunit
C:Keywords: deoxyribonucleotide biosynthesis; oxidoreductase

Alignment Scores:
Pred. NO.: 2.73e-108 Length: 319
Score: 1280.00 Matches: 239
Percent Similarity: 85.31% Conservative: 34
Best Local Similarity: 74.69% Mismatches: 45
Query Match: 69.34% Indels: 2
DB: 2 Gaps: 2

US-10-698-228-12 (1-1053) x T30782 (1-319)

QY 94 GAGCCACTCCTCAAGAAAGAGTTCTCGCGGTTTGTCTCATCTTTCCAATCCAGTACCTGAT 153
Db 2 GluProIleLeuAlaProAsnProAsnArgPheValIlePheProIleGlnTyrHisAsp 21
QY 154 ATTTGGAAATATGATATAACAGGACAGGCTTCTTCGACAGCAGCAGAGAGGTTGACTTA 213
Db 22 IleTrpAsnMetTyrLysLysAlaGluAlaSerPheTrpThrValGluGluValAspIle 41
QY 214 TCAAGAGATCTCCTCACTCGAAACAAGCTTAAACAGCAGATGAGAGTACTTCATCTCTCAC 273
Db 42 SerLysAspIleAsnAspTrpAsnLysLeuThrProAspGluLysTyrPheLysHis 61
QY 274 ATCTTAGCCCTTTTTCAGCCAGTGTGAATTTAAATGAAATTTTGGTGGAGCGCTTT 333
Db 62 ValLeuAlaPhePheAlaAlaSerAspGlyLysValAsnGluAsnLeuAlaGluArgPhe 81
QY 334 AGTCAGAGATGTCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAAATTTCTCATCAG 393
Db 82 CysThrGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnMetAlaIleGlu 101
QY 394 AATGTTCACTCAGAGATGTACAGTTTCTGATAGACACTTACATCAGAGATCCCAAGAAA 453
Db 102 AsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrValLysAspSerAsnGlu 121
QY 454 AGGGAATTTTATTAAATGCAATTGAAACCATGCCCTATGTTAAAGAAAAAGACAGATTGG 513
Db 122 LysAsnTyrLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAspTrp 141
QY 514 GCCTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 573
Db 142 AlaGlnLysTrpIleHisAsp---SerAlaGlyTyrGlyGluArgLeuIleAlaPheAla 160
QY 574 GCTGTAGAAGAGATTTTCTCTCAGGATCTTTGCTGCTATATTTCTGGCTAAAGAGAGA 633
Db 161 AlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArg 180
QY 634 GGTCTTATGCCAGACTCAGTTTTCCTCAATGAACTCATCAGCAGATGAAAGGACTTCAC 693
Db 181 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHis 200
QY 694 TGTGACTTTCTGCTGCTGATGTTTCCAAATCTTCAATTAATAGCCTTCAGAGAAAGGGTCT 753
Db 201 CysAspPheAlaCysLeuMetPheLysHisLeuLeuHisProProSerGluGluThrVal 220
QY 754 AGGAGATCATTTGTTGATGCTGCTCAAAATTTGAGCAGGAGTTTAAACAGAGCTTGCA 813
Db 221 ArgSerIleIleThrAspAlaValSerIleGluGlnGluPheLeuThrAlaAlaLeuPro 240
QY 814 GTTGGCTCACTGGAATGATTCATTTTTCATGAAACAGTACATTTGAGTTTGTAGCTGAC 873
Db 241 ValLysLeuIleGlyMetAsnCysGluMetLysThrTyrIleGluPheValAlaAsp 260
QY 874 AGATTACTTGTGGAACCTTGATTTCTCAAGATTTTTCAGGCGAGAAAAATCCTTTGATTTT 933
Db 261 ArgLeuIleSerGluLeuGlyPheLysLysIleTyrAsnValThrAsnProPheAspPhe 280
QY 934 ATGGAACAACTTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAGAGTAT 993
Db 281 MetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGluTyr 300
QY 994 CAGCGTTTTCAGTTATGCGAGAAACCCACAGATTAACCTTTCACCTTGGATGCGAGATTT 1053
Db 301 GlnLysMetGlyValMetSerGln---GluAspAsnHisPheSerLeuAspValAspPhe 319
RESULT 7
B72154
E4L protein - variola minor virus (strain Garcia-1966)
C:Species: variola minor virus
C:Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 09-Jul-2004
C:Accession: B72154
R:Shchelkunov, S.N.; Totmenin, A.V.; Gutorov, V.V.; Safronov, P.F.; Massung, R.F.; Lopar
submitted to GenBank, March 1998
A:Description: Analysis of the complete coding sequence of DNA of alastrim variola minor

A:Reference number: A72150
A:Accession: B72154
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-319 <SHC>
A:Cross-references: UNIPROT:Q89087; GB:Y16780; NID:G5803555; PIDN:CAB54628.1; PID:G5803555
A:Experimental source: strain Garcia-1966
C:Genetics:
A:Gene: E4L
C:Superfamily: ribonucleoside reductase small subunit

Alignment Scores: 2.73e-108 Length: 319
Pred. No.: 1280.00 Matches: 240
Score: 85.31% Conservative: 33
Percent Similarity: 75.00% Mismatches: 45
Best Local Similarity: 69.34% Indels: 2
Query Match: 2
DB: Gaps: 2

US-10-698-228-12 (1-1053) x B72154 (1-319)

QY 94 GAGCCACTCTAAGAAGAGTTTCGCCGGTGTTCATCTTTCAAATCCAGTACCTCGAT 153
Db 2 GluProileuAlaIysAsnProAsnArgPheValilePheProileGlnTyHisasp 21
QY 154 ATTTCGAAATGATAAACAGGCACAGGCTTCCTCTGGACAGACAGAGGTGGACTTA 213
Db 22 IleTrpAsnMetTyrlsYsAlaGluAlaSerPheTrpThrValGluGluValAspIle 41
QY 214 TCAAAGGATCTCCCTCACTCGGAACAAGCTTAAAGCAGATGAGAAGTACTTCATCTCTCAC 273
Db 42 SerLysAspIleAsnAspTrpAsnLysLeuThrProAspGluLysTyrlleLysHis 61
QY 274 ATCTTAGCCTTTTTCGACGACGATGGAATTTGTAATCAAAATTTGGTGGACGCTTT 333
Db 62 ValLeuAlaPhePheAlaAlaSerAspGlyileValAsnGluAsnLeuAlaGluA-gPhe 81
QY 334 AGTCAGAGGTGCAGGTTCCAGAGGTCGCTGTTCATGGCTTCAAAATTCATCTCGAG 393
Db 82 CysIleGluValGlnIleThrGluAlaA-gCysPheTyrlGlyPheGlnMetAlaIleGlu 101
QY 394 AATGTTCACCTCAGAGATGTACAGTTTGCTGATAGACACTTACATCAGATGCCAAGAAA 453
Db 102 AsnIleHisSerGluMetTyrlSerLeuLeuIleAspThrTyrlValLysAspSerAsnGlu 121
QY 454 AGGAAATTTTATTAATGCAATTGAAACCATGCCCTATGTTAAGAAAAAAGCAGATGG 513
Db 122 LysAsnTyrlleuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAspTrp 141
QY 514 GCCTTCGGATGATACAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCTTTGCT 573
Db 142 AlaGlnLysTrpIleHisasp---SerAlaGlyTyrlGlyGluArgLeuIleAlaPheAla 160
QY 574 GCTGTGAAGAGGATTTTCTTCAGGATCTTTTGCTGCTATATTCCTGGCTAAAGAAAGA 633
Db 161 AlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArg 180
QY 634 GGTCTTATGCCAGGACTCACTTTTCCATGAATCATCAGCAGAGATGAAGACTTCAC 693
Db 181 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHis 200
QY 694 TGTGACTTTGCTGCTGATGTTTCCAATCTAGTAAATTAAGCTTCAGAAAGAAAGGTC 753
Db 201 CysAspPheAlaCysLeuMetPheLysHisLeuLeuTyrlProSerGluGluThrVal 220
QY 754 AGCGAGATCATTTGATGCTGTCAAAATTCAGCAGGAGTTTTTAAACAGAGCCTTGCCA 813
Db 221 ArgSerIleIleThrAspAlaValSerIleGluGlnGluPheLeuThrValAlaLeuPro 240
QY 814 GTTGGCCTCATTCGAATGAATTCATTTGATGAACACAGTACATTGAGTTGTAGCTGAC 873
Db 241 ValLysLeuIleGlyMetAsnCysGluMetLysThrTyrlleGluPheValAlaAsp 260

Qy	874	AGATTACTGTGGAACTTGATCTCTCAAGGCTTTTTCAGGAGAGAAATCCCTTTGATTTT	933
Db	261	ArgLeuIleSerGluLeuGlyPheTyrSlyIleTyrAsnValThrAsnProPheAspPhe	280
Qy	934	ATGGAACATTTCTTTAGAACGAGAAACAAATTTCTTTGAGAACGAGTTTCAGAGTAT	993
Db	281	MetGluAsnIleSerLeuGluGlyIleThrAsnPhePheGluLysArgValGlyGluTyr	300
Qy	994	CAGCGTTTTTGCACTTATGCGAGAAACACAGATAACGCTCTTACCTTGGATGCAGATTTT	1053
Db	301	GlnLysMetGlyValMetSerGln---GluAspAsnHisPheSerLeuAspValAspPhe	319
RESULT 8			
RDVZVV			
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - vaccinia virus			
N:Alternate names: F4L protein			
C:Species: vaccinia virus			
C:Date: 30-Jun-1989 #sequence revision 30-Jun-1989 #text_change 09-Jul-2004			
C:Accession: A29892; 142506; F36213			
R:Slabaugh, M.; Roseman, N.; Davis, R.; Mathews, C.			
J. Virol. 62, 519-527, 1988			
A:Title: Vaccinia virus-encoded ribonucleotide reductase: sequence conservation of the			
A:Reference number: A29892; MUID:88091062; PMID:2826813			
A:Accession: A29892			
A:Molecule type: DNA			
A:Residues: 1-212, 'Y', 214-319 <G>			
A:Cross-references: GB:M5027; NID:G335317; PIDN:AAA48018.1; PID:G335366			
A:Experimental source: strain Copenhagen			
R:Goebel, S.J.; Johnson, G.P.; Perkus, M.E.; Davis, S.W.; Winslow, J.P.; Paoletti, E.			
Virology 179, 247-266, 1990			
A:Title: The complete DNA sequence of vaccinia virus.			
A:Reference number: A42531; MUID:91021027; PMID:2219722			
A:Contents: annotation; possible protein-coding frames			
A:Note: neither amino acid nor nucleotide sequence is given			
R:Roseman, N.A.; Slabaugh, M.B.			
Virology 178, 410-418, 1990			
A:Title: The vaccinia virus HindIII F fragment: nucleotide sequence of the left 6.2 kb.			
A:Reference number: A36213; MUID:91020979; PMID:2219701			
A:Accession: F36213			
A>Status: preliminary			
A:Molecule type: DNA			
A:Residues: 1-319 <ROS>			
A:Cross-references: EMBL:M34368; NID:G335618; PIDN:AAA48244.1; PID:G335624			
A:Experimental source: strain WR			
C:Function:			
A:Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleoside			
C:Superfamily: ribonucleoside reductase small subunit			
C:Keywords: DNA replication; iron; metalloprotein; oxidoreductase			
F:70,101,104,163,197,200/Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, His) #sta			
F:108/Active site: Tyr (stable tyrosyl radical) #status predicted			
Alignment Scores:			
Pred. No.:		3,36e-108	Length: 319
Score:		1279.00	Matches: 239
Percent Similarity:		85.31%	Conservative: 34
Best Local Similarity:		74.63%	Mismatches: 45
Query Match:		69.28%	Indels: 2
DB:		1	Gaps: 2
US-10-698-228-12 (1-1053) x RDVZVV (1-319)			
Qy	94	GAGCCACTCTTAAGAAAGATGTCGCCGGTTTGTCATCTTTCCAATCCAGTACCTGAT	153
Db	2	GluProIleLeuAlaProAsnArgPheValIlePheProIleGlnTyrAsp	21

218 IleAlaPheAlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrp 237

622 CTAAGAAGAGAGAGGCTTTATGCGAGACTCACTTTTCCAAATGAACCTCATCAGCAGAGAT 681

238 LeuLysLysArgGlyLeuMetProGlyLeuThrHisSerAsnGluLeuIleSerArgAsp 257

682 GAAGGACTTCACTGTGACTTTGCTTGCCTGTGATGTTCCCAATACTATTAGTAAATTAAGCCTTCA 741

258 GluGlyLeuHisArgAspPheAlaCysLeuLeuTyrSerLysLeuGlnLysLysLeuThr 277

742 GAAGAAGGGTCAGGAGCATCATTTGATGCTGTGCTCAAAATTCAGCAGGAGTTTAAACA 801

278 GlnGlnArgIleTyrAspIleLysAspAlaValAlaIleGluGlnGluPheLeuThr 297

802 GAAGCCTTGCAGTTGGCCTCATTTGGAATGAATTCATTTTGATGAACAGTACATTGAG 861

298 GluAlaLeuProValaspMetIleGlyMetAsnCysArgLeuMetSerGlnTyrIleGlu 317

862 TTTGTAGCTGACAGATTACTTGTGAACTTGAACTTCAAAGTTTTTCAGGCAGAAAT 921

318 PheValAlaAspHisLeuLeuValGluLeuGlyCysAspLysLeuTyrLysSerLysAsn 337

922 CCTTTTCAGTTTATGGAACAACTTTCTTTAGAGGAGAAACAAATTTCTTTTGAGAAACGA 981

338 ProPheaspPheMetGluAsnIleSerIleAspGlyLysThrAsnPhePheGluLysArg 357

982 GTTTCAGAGTATCAGCGTTTTTCAGTTATGGCAGAAACCCACAGATAACGCTTTCACCTTG 1041

358 ValSerGluTyrGlnArgProGlyValMetValAsnGluAlaGluArgGlnPheAspLeu 377

1042 GATGCGAGATTTT 1053

378 GluAlaAspPhe 381

RESULT 14

A26916

ribonucleoside-diphosphate reductase (EC 1.1.7.4.1) small chain - yeast (Saccharomyces cerevisiae)

N;Alternate names: protein J1271; protein YJL026w; ribonucleotide reductase small chain

C;Species: Saccharomyces cerevisiae

C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 09-Jul-2004

R;Elledge, S.J.; Davis, R.W.

Mol. Cell. Biol. 7, 2783-2793, 1987

A;Title: Identification and isolation of the gene encoding the small subunit of ribonucleotide reductase

A;Reference number: A26916; MUID:88038817; PMID:3313004

A;Accession: A26916

A;Molecule type: DNA

A;Residues: 1-399 <ELL>

A;Cross-references: UNIPROT:P09938; EMBL:M17221

A;Note: the authors translated the codon GAA for residue 101 as Lys, GCT for residue 102 as Ser, GCT for residue 107 as Phe, and GAA for residue 108 as Trp

A;Note: the sequence shown follows the authors' translation

R;Hurd, H.K.; Roberts, C.W.; Roberts, J.W.

Mol. Cell. Biol. 7, 3673-3677, 1987

A;Title: Identification of the gene for the yeast ribonucleotide reductase small subunit

A;Reference number: S07605; MUID:88065506; PMID:3316984

A;Accession: S07605

A;Molecule type: DNA

A;Residues: 1-399 <HUR>

A;Cross-references: EMBL:M17789; NID:gi172449; PIDN:AAA34988.1; PID:gl172450

R;Pohl, T.M.; Aljinovic, G.

submitted to the Protein Sequence Database, September 1995

A;Reference number: S56793

A;Accession: S56798

A;Molecule type: DNA

A;Residues: 1-399 <TOV>

A;Cross-references: EMBL:Z49301; NID:gi008141; PIDN:CAA89317.1; PID:gi008142; GSPDB:GN0001

C;Genetics:

A;Gene: SGD:RNR2; MIPS:YJL026w

A;Cross-references: SGD:S0003563; MIPS:YJL026w

A;Map position: 10L

C;Function:

A;Description: oxidoreductase; pyrimidine deoxynucleotide metabolism

C:Superfamily: ribonucleoside reductase small subunit
C:Keywords: iron; oxidoreductase; pyrimidine deoxynucleotide metabolism
F:145,176,179,239,273,276/Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, His) #ad
F:183/Active site: Tyr (stable tyrosyl radical) #status predicted

Alignment Scores:
Pred. No.: 3,286-98 Length: 399
Score: 1169.50 Matches: 221
Percent Similarity: 81.07% Conservative: 53
Best Local Similarity: 65.36% Mismatches: 58
Query Match: 63.35% Indels: 6
DB: 1 Gaps: 3

US-10-698-228-12 (1-1053) x A26916 (1-399)

QY 54 ATCTTCAGACCAACCAAGTGAATAAAGTCAATGAAGCCACTCTTAAGAAGAG 113
DB 64 VAHISARGHISLysLeuLys-GluMetGlu---LysGluGluProLeuLeuAsnGluAs 82
QY 114 TTCTCGCGGTTGTTCATCTTCCCAATCCAGTACCTGATATTGGAATAATGTATAACA 173
DB 82 physGluArgThrValLeuPheProIleLysTyrHisGluIleTrpGlnAlaTyrLysAr 102
QY 174 GGCACAGCTTCTCTTCGACAGCAGAGGTGTGACTTATCAAGGATCTCCCTCAGTG 233
DB 102 galAGluAlaSerPheTrpThrAlaGluIleAspLeuSerLysAspIleHisAspTr 122
QY 234 G---NACAGCTTAAGCAGATGAGAGTACTTCATCTTCATCTTACATCTTAGCCTTTTTC 290
DB 122 pAsnAsnArgMetAsnGluAsnGluArgPheIleSerArgValLeuAlaPheAl 142
QY 291 AGCCAGTATGAAATTTGTAATGAAATTTGTGGAGCGCTTTAGTCAGAGGTGCAGGT 350
DB 142 alaserAspGlyIleValAsnGluAsnLeuValGluAsnPheSerThrGluValGlnI 162
QY 351 TCCAGAGCTCGCTGTTCTTATGGCTTTTCAAAATTTCTCATCGAGAAATGTTCACTCAGAGAT 410
DB 162 eProGluAlaLysSerPheTyrGlyPheGlnIleMetIleGluAsnIleHisSerGluTh 182
QY 411 GTACAGTTGCTGTAGACACTTACATCAGACATCCCAAGAAAGGAAATTTTATTAA 470
DB 182 rTyrSerLeuLeuIleAspThrTyrIleLysAspProLysGluSerGluPheLeuPheAs 202
QY 471 TGCATTGAACCATGCCCTATGTTAAGAAAGCAGATTTGGGCTTCGGATGATAGC 530
DB 202 nAlaIleHisThrIleProGluIleGlyGluLysAlaGluTrpAlaLeuArgTrpIleG 222
QY 531 AGATAGAAATCTACTTTTGGGAAAGAGTGTGGCTTTGTCTGTGTAGAAGGAGTTT 590
DB 222 nAspAlaAspAlaLeuPheGlyGluArgLeuValAlaPheAlaSerIleGluGlyValPh 242
QY 591 CTCTCAGGATCTTTTGTGCTATATTCTGGTAAAGAGAGAGGTCTTATGCCAGGACT 650
DB 242 ePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArgGlyMetMetProGlyLe 262
QY 651 CACTTTTTCATGAACTCATCAGCAGATGAGGACTTCACTGTGACTTTCGCTTGCCT 710
DB 262 uThrPheSerAsnGluLeuIleCysArgAspGluGlyLeuHisThrAspPheAlaCysLe 282
QY 711 GATGTTCCAAATCTAGTAAATAAGCTTTCAGAAAGAGGTCAGGAGATCATTTGTGA 770
DB 282 uLeuPheAlaHisLeuLysAsnLysProAspProAlaIleValGluLysIleValThrGl 302
QY 771 TGCTGTCAAAATTCAGCAGGAGTTTTCACAGAGCTTTGCCAGTTGGCTCATTTGGAAT 830
DB 302 uAlaValGluIleGluGlnArgTyrPheLeuAspAlaLeuProValAlaLeuGlyLe 322
QY 831 GAATTGCAATTTGTGAACACAGTACATCAGTGTGTGTAGCTGACAGATTACTTGTGAAC 890
DB 322 tAsnAlaAspLeuMetAsnGlnTyrValGluPheValAlaAspArgLeuLeuValAlaPh 342
QY 891 TGGATTCTCAAGGTTTTCAGGCGAATAATCTTTGATTTTATGGAACAATCTTTT 950

Db 342 eGlyAsnLysLysTyrTyrLysValGluAsnProPheAspPheMetGluAsnIleSerLe 362
QY 951 AGAAGGAAACAAATTTCTTTGAGAAACAGAGTTTCAGAGTATCAGCGCTTTGCAGATTAT 1010
Db 362 uAlaGlyLysThrAsnPhePheGluLysArgValSerAspTyrGlnLysAlaGlyValMe 382
QY 1011 GGCAGAAACCCACAGAT-----AACGTCTTCCACCTTGGATGCGAGATT 1053
Db 382 tSerLysSerThrLysGlnGluAlaGlyAlaPheThrPheAsnGluAspPhe 399
RESULT 15
T39992
ribonucleoside-diphosphate reductase small chain - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T39992
R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Whithead, S.; Chillingworth, T.; Churcher, S.;
submitted to the EMBL Data Library, July 1998
A:Reference number: Z21897
A:Accession: T39992
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-391 <LYN>
A:Cross-references: UNIPROT:P36603; EMBL:AL031158; PIDN:CAA20100.1; GSPDB:GN00067; SPDB:
A:Experimental source: strain 972h-; cosmid c25D12
C:Genetics:
A:Gene: SPDB:SPBC25D12.04
A:Map position: 2
C:Superfamily: ribonucleoside reductase small subunit

Alignment Scores:
Pred. No.: 1,756-97 Length: 391
Score: 1161.50 Matches: 227
Percent Similarity: 78.80% Conservative: 48
Best Local Similarity: 65.04% Mismatches: 61
Query Match: 62.92% Indels: 13
DB: 2 Gaps: 4

US-10-698-228-12 (1-1053) x T39992 (1-391)

QY 37 GATCAGATGAGAGATCATCTTCAGAC-----ACCAACGAAAGTGAATAAAG 84
DB 46 GluGluLysGlnGluGlyAspTyrTyrLeuGlyLysLysGluAspGluLeu--- 64
QY 85 TCAATGAAGAGCCACTCTTAAAGAGAGTCTCGCGGTTTGTTCATCTTCCAAATCCAG 144
DB 65 -----AspGluValValLeuArgProAsnProHisArgPheValLeuPheProIleLys 82
QY 145 TACCCTGATATTGGAAATGTATAACAGCCACAGGCTTCCTTCGACAGCAGAGAG 204
DB 83 TyrHisGluIleTrpGlnPheTyrLysAlaGluAlaSerPheTrpThrAlaGluGlu 102
QY 205 GTTGACTTATCAAGGATCTCCCTCACTGG---AACCAAGCTTAAAGCAGATGAGAGATGAC 261
DB 103 IleAspLeuSerLysAspLeuValAspTrpAspAsnLysLeuAsnAlaAspGluArgTyr 122
QY 262 TTCATCTCTCAGATCTAGCTTTTTCGACCCAGTGTGAATGTAAATGAAATTTG 321
DB 123 PheIleSerThrValLeuAlaTyrPheAlaSerAspGlyIleValAsnGluAsnLeu 142
QY 322 GTGGAGCGCTTTAGTCAGGAGGTTCAGGCTTCAGAGGCTCGCTGTTTCTATGGCTTCAA 381
DB 143 LeuGluArgPheSerSerGluValGlnIleProGluAlaArgCysValTyrGlyPheGln 162
QY 382 ATTCTCATTCAGAAATGTTCACTCAGAGATGTACAGTTTGTGTAGATAGACACTTACATCAGA 441
DB 163 IleMetIleGluAsnIleHisSerGluThrTyrSerLeuLeuLeuAspThrTyrIleArg 182
QY 442 GATCCCAAGAAAGGAAATTTTATTAAATGCAATGAAACCATGCCCTATGTTAAGAAA 501
DB 183 GluProLysGluLysGlnArgHisPheAspAlaIleLeuThrMetGlySerIleLysAla 202
QY 502 AAAGCAGATTGGCGCTTTCGATGATGATGAGATAAATCTACTTTTGGGGAAGAGATG 561

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 30, 2005, 03:28:20 ; Search time 76 Seconds
(without alignments)
14189.984 Million cell updates/sec

Title: US-10-698-228-12

Perfect score: 1846

Sequence: 1 atggcgaccggaaagcc.....tcaccttgatcgagatttt 1053

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US10698228/runat_26102005_100608_3981/app.query.fasta.1.2446
-DB=UniProt_03 -QFMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -ENDS=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=PCT -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228 @Cgn 1.1.149 @runat_26102005_100608_3981 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORE=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt_03.*

1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	98.6	351	2 Q7LG56	Q7LG56 homo sapien
2	1817	98.4	351	2 Q9NUW3	Q9NUW3 homo sapien
3	1725	93.4	351	2 Q6PEE3	Q6PEE3 mus musculus
4	1534	83.1	297	2 Q6A141	Q6A141 homo sapien
5	1514	82.0	299	2 Q7SPY9	Q7SPY9 homo sapien
6	1474	79.8	285	2 Q86YE3	Q86YE3 homo sapien
7	1454	78.8	389	1 R1R2 HUMAN	P11157 mus musculus
8	1443	78.2	390	1 R1R2 MOUSE	P11157 mus musculus
9	1439	78.0	386	2 Q68EE1	Q68EE1 xenopus tro
10	1431	77.5	386	2 Q801Q4	Q801Q4 xenopus lae
11	1427	77.3	386	2 Q8AVY2	Q8AVY2 xenopus lae
12	1425.5	77.2	386	1 R1R2 BRARE	P79733 brachydanio
13	1421	77.0	386	1 R1R2 MESAU	Q60561 mesocricetu
14	1420.5	77.0	386	2 Q6D144	Q6D144 brachydanio
15	1399.5	75.8	378	2 Q6P876	Q6P876 xenopus tro
16	1385	75.0	406	2 Q6IP47	Q6IP47 xenopus lae

17	1382.5	74.9	388	2	Q7PF28	Q7PF28 anopheles g
18	1382.5	74.9	426	2	Q7QIF4	Q7QIF4 anopheles g
19	1370	74.2	349	2	Q7ZYW0	Q7ZYW0 brachydanio
20	1358	73.6	397	2	Q95VP8	Q95VP8 aedes aegypt
21	1349.5	73.1	399	2	Q27124	Q27124 urechis cau
22	1348	73.0	393	1	R1R2 DROME	P48592 drosophila
23	1335	72.3	398	2	Q9XYN8	Q9XYN8 aedes albop
24	1331	72.1	384	1	R1R2 SPISO	P07201 spigula sol
25	1284	69.5	403	2	Q6CFU6	Q6CFU6 yarrowia li
26	1280	69.3	319	2	Q57175	Q57175 vaccinia vi
27	1280	69.3	319	2	Q76Q46	Q76Q46 variola min
28	1280	69.3	319	2	Q89087	Q89087 variola vir
29	1279	69.3	319	1	R1R2 VACCV	P11158 vaccinia vi
30	1279	69.3	319	2	Q76ZK1	Q76ZK1 vaccinia vi
31	1277	69.2	319	2	Q6R209	Q6R209 rabbitpox v
32	1276	69.1	319	1	R1R2 VACCC	P20493 vaccinia vi
33	1276	69.1	319	2	Q8JLH6	Q8JLH6 ectromelia
34	1276	69.1	319	2	Q89559	Q89559 variola vir
35	1274	69.0	333	2	Q8QN14	Q8QN14 cowpox viru
36	1273	69.0	319	1	R1R2 VACCP	P29883 vaccinia vi
37	1271	68.9	319	2	Q8V544	Q8V544 monkeypox v
38	1270	68.8	319	2	Q87632	Q87632 cowpox viru
39	1268	68.7	319	2	Q8V2X8	Q8V2X8 camelpox vi
40	1268	68.7	319	2	Q77525	Q77525 camelpox vi
41	1267	68.6	319	2	Q9JPE9	Q9JPE9 vaccinia vi
42	1266	68.6	319	1	R1R2 VARV	P31799 variola vir
43	1264.5	68.5	410	1	R1R2 NEUCR	Q9C167 neurospora
44	1260	68.3	333	2	Q9PXR4	Q9PXR4 variola vir
45	1258.5	68.2	348	2	Q9YMK7	Q9YMK7 lymantria d

ALIGNMENTS

RESULT 1

Q7LG56 PRELIMINARY; PRT; 351 AA.
ID Q7LG56 Q9NPD6; Q9NTD8;
AC Q7LG56; Q9NPD6; Q9NTD8; 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 28, Last annotation update)
DE Ribonucleotide reductase (Hypothetical protein DKEZp761E1312) (P53-
DE Inducible ribonucleotide reductase small subunit 2).
GN Name=p53R2; Synonyms=DKEZp761E1312;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20179179; PubMed=10716435; DOI=10.1038/35003506;
RA Tanaka H., Arakawa H., Yamaguchi T., Shiraishi K., Fukuda S.,
RA Matsui K., Takei Y., Nakamura Y.,
RT "A ribonucleotide reductase gene involved in a p53-dependent cell-
RT cycle checkpoint for DNA damage.";
RL Nature 404:42-49(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Amalgam;
RG The German cDNA Consortium;
RA Ansoorge W., Krieger S., Regiert T., Rittmuller C., Schwager B.,
RA Mewes H.W., Weill B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Ugai H., Yokoyama K.K.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB036532; BAA92493.1; JOINED.
DR EMBL; AB036524; BAA92493.1; JOINED.
DR EMBL; AB036525; BAA92493.1; JOINED.
DR EMBL; AB036526; BAA92493.1; JOINED.
DR EMBL; AB036527; BAA92493.1; JOINED.
DR EMBL; AB036528; BAA92493.1; JOINED.
DR EMBL; AB036529; BAA92493.1; JOINED.

DR GO: GO:0004748; F:ribonucleoside-diphosphate reductase activity; IEA.
 DR GO: GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.
 DR InterPro: IPR009078; Ferritin/RR like.
 DR InterPro: IPR000358; Ribonucleotide reductase.
 DR Pfam: PF00268; Ribonucleotide reductase.
 DR PROSITE: PS00368; RIBOSOMAL SMALL; 1.
 SQ SEQUENCE 351 AA; 40704 MW; 6D009B2D59E9A323 CRC64;

Alignment Scores:

Pred. No.: 7,15e-151 Length: 351
 Score: 1917.00 Matches: 350
 Percent Similarity: 100.00% Conservatives: 1
 Best Local Similarity: 99.72% Mismatches: 0
 Query Match: 98.43% Indels: 0
 DB: 2 Gaps: 0

US-10-698-228-12 (1-1053) x Q9NUW3 (1-351)

QY 1 ATGGGCGACCCGGAAGCGCGGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
 Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
 QY 61 GACACACGAAGTGAATAAGTCAATGAAGAGCCACTCTAAGAAAGAGTCTCCG 120
 Db 21 AspThrAsnGluSerGluLeuLeuSerAsnGluGluProLeuLeuArgLysSerArg 40
 QY 121 CGGTTGTCATCTTCCATCCAGTACCTGATATTTGGAATGTAATAACAGGACAG 180
 Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
 QY 181 GCTTCTCTCTGACAGCAGAGAGGTGCTACTTATCAAGGATCTCCTCAGTGAACAAG 240
 Db 61 AlaSerPheThrPheAlaGluValAspLeuSerLysAspLeuProIleThrAsnLys 80
 QY 241 CTTAAAGCAGATGAGAAGTACTCTCTACATCTTACCTTTTGGACCCAGTGTAT 300
 Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisLeuAlaPheAlaAspAsp 100
 QY 301 GGAATTGTAATGAATTTGTGAGCGCTTTAGTACGAGGTGAGGTCAGAGGCT 360
 Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValProGluAla 120
 QY 361 CGCTGTTCTAGGCTTCAATCTCTCATCGAATGTTCTCAGAGATGTTACAGTTTG 420
 Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGA 480
 Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 QY 481 ACCATGCGCTATGTTAAGAAAGACAGATTGGCGCTTGCATGATGATGATGAGAAA 540
 Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 QY 541 TCTACTTTGGGAAGAGTGTGGCTTGTCTGTAGAGAGGTTTCTTCTCAGA 600
 Db 181 SerThrPheGlyGluArgValAlaPheAlaValGluGluValPhePheSerGly 200
 QY 601 TCTTTTGTCTATATCTGGCTTAAGAAAGAGAGGTCTTATCCAGGAGTCACTTTTCC 660
 Db 201 SerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 QY 661 AATGAACATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGTGCTGTGTTCCAA 720
 Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 QY 721 TACTTAGTAATACCTTCAAGAAAGGTCAGGAGATCATTTGATGCTGTCAA 780
 Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260
 QY 781 ATTGAGCAGGAGCTTTTAAACAGACCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 840
 Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyValAsnCysIle 280

QY 841 TTGATGAAAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTGTGGAACTTGGATTCTCA 900
 Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 QY 901 AAGTTTTCAGGACGAGAAATCTTTGATTTTATGAAACATTTCTTTAGAAAGGAAAA 960
 Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 QY 961 ACAATTTCTTTCAGAAACGAGTTTCAGACTATCAGGTTTCAGTTTTCAGTTTTCAGTATG 1020
 Db 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 QY 1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053
 Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 3

Q6PEE3 PRELIMINARY; PRT; 351 AA.
 AC O6PEE3; 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, last annotation update)
 DE Ribonucleotide reductase M2 B (Tp53 inducible).
 GN Name=Rm2b;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuetz G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ussidi T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Rosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalek U., Small U., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC058103; AAH58103.1; -
 DR GO; GO:0004748; Ribonucleoside-diphosphate reductase activity; IEA.
 DR GO; GO:0009263; P:deoxyribonucleotide biosynthesis; IEA.
 DR InterPro; IPR009078; Ferritin/RR like.
 DR InterPro; IPR000358; Ribonucleotide reductase.
 DR Pfam; PF00268; Ribonucleotide reductase.
 DR PROSITE; PS00368; RIBOSOMAL SMALL; 1.
 SQ SEQUENCE 351 AA; 40803 MW; 4E1259233C9CC8A9 CRC64;

Alignment Scores:

Pred. No.: 8,75e-143 Length: 351
 Score: 1725.00 Matches: 328
 Percent Similarity: 97.44% Conservatives: 14
 Best Local Similarity: 93.45% Mismatches: 9

Query Match:	93.45%	Indels:	0
DB:	2	Gaps:	0
US-10-698-228-12 (1-1053) x Q6PEB3 (1-351)			
QY	1	ATGGCGCACCGGAAAGCGCGGCGGTGGATCAGGATCAGATCATCTTCA	60
DB	1	MetGlyAspProGluArgProGluAlaAaArgProGluLysGlyGluInLeuCysSer	20
QY	61	GACACCAAGCAAGTAAATCAATGAAGACCACTCTTAAAGAAAGATCTTCGC	120
DB	21	GluThrGluGluAsnValValArgSerAsnGluProLeuLeuArgLysSerSerArg	40
QY	121	CGTTTGTCTATCTTCCATCCAGTACCCTGATATTTCGAAAATGTATAAAGGACAG	180
DB	41	ArgPheValIlePheProIleGlnTyrProAspIleTrpArgMetTyrLysGlnAlaGln	60
QY	181	GTTCTCTTCGACACAGACAGAGGTGACTTATCAAGAGTCTCCCTCACTGGAACAAG	240
DB	61	AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys	80
QY	241	CTTAAGCAGATGAGAAGTACTTCTCATCTCTACATCTTAGCCTTTTTCAGCCAGTGAT	300
DB	81	LeuLysSerAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAaSerAsp	100
QY	301	GGAAATTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT	360
DB	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla	120
QY	361	CGCTGTTTCTATGGCTTCAAAATCTCATCGAATGTTCACTCAGAGATGTACAGTTTG	420
DB	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140
QY	421	CTGATAGACACTTATCATCAGATCCCAAGAAAGGGAATTTTATTAATGCAATTCGAA	480
DB	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160
QY	481	ACCATGCCCTATGTTAAGAAAAGCAGATTGGGCTTCGATGGATGATGACAGATAGAAA	540
DB	161	ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys	180
QY	541	TCTACTTTTGGGAAAGAGTGTGGCTTCTGCTGTAGAGAGTCTTCTTCTCAGGA	600
DB	181	SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyIlePhePheSerGly	200
QY	601	TCTTTTGTGCTATATCTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACATTTTCC	660
DB	201	SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer	220
QY	661	AATGAATCATCAGCAGATGAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA	720
DB	221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln	240
QY	721	TACTTAGTAAATAGCCTTTCAGAAAGAGGTTCAGGAGATCATTTGTGATGCTGTCAAA	780
DB	241	TyrLeuValAsnLysProSerGluAspArgValArgGluIleIleAlaAspAlaValGln	260
QY	781	ATTGACGAGGCTTTTAAACAGAGCCTTGCAGTTGGCCTCATTCGGAATCAATTCATT	840
DB	261	IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysVal	280
QY	841	TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTTCA	900
DB	281	LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuGlyGlyLeuGlyPheSer	300
QY	901	AAGGTTTTTCAGCGAGAAAATCTTTTGAATTTATGGAACCAATTTCTTTAGAGGAAA	960
DB	301	LysIlePheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys	320
QY	961	ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTGTGACGTTATGGCAGAAC	1020
DB	321	ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr	340

QY	1021	ACAGATAACGCTTCTTACCTTGGATCAGATTTT	1053
DB	341	ThrAspAsnValPheThrLeuAspAlaAspPhe	351

RESULT 4

Q6AI41

Q6AI41

PRELIMINARY;

PRT;

297 AA.

AC

O6AI41;

DT

25-OCT-2004

(TrEMBLrel. 28, Created)

DT

25-OCT-2004

(TrEMBLrel. 28, Last sequence update)

DT

25-OCT-2004

(TrEMBLrel. 28, Last annotation update)

DE

Hypothetical protein DKFZp686M05248.

GN

Name=DKFZp686M05248;

OS

Homo sapiens (Human).

OC

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX

NCBI_TaxID=9606;

RN

[1]

RP

SEQUENCE FROM N.A.

RC

TISSUE=Salivary gland;

RG

The German cDNA Consortium;

RA

Ottewaelder B., Obermaier B., Deutschenbaur S., Schaipp A.,

RA

Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;

RL

Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.

DR

EMBL; CR627376; CAH10473.1;

DR

GO; GO:0004748; Fibrinucleoside-diphosphate reductase activity; IEA.

DR

GO; GO:0009186; P-deoxyribonucleoside diphosphate metabolism; IEA.

DR

InterPro; IPR009078; Ferritin/RR like.

DR

InterPro; IPR000358; Ribonuc_redctse.

DR

Pfam; PF00268; Ribonuc_red_sm; 1.

DR

PROSITE; PS00368; RIBORED_SMALL; 1.

KW

Hypothetical protein.

SQ

SEQUENCE 297 AA; 34498 MW; 2BE522P2CCI58A02 CRC64;

Alignment Scores:

Pred. No.:

5.28e-126

Length:

297

Score:

1534.00

Matches:

296

Percent Similarity:

99.66%

Conservative:

0

Best Local Similarity:

99.66%

Mismatches:

1

Query Match:

83.10%

Indels:

0

Gaps:

0

US-10-698-228-12 (1-1053) x Q6AI41 (1-297)

QY

163

ATGTATAACAGCAGCAGGCTTCTTCGACAGCAGAGGTTGACTTATCAAGGAT

222

DB

1

MetTyrLysGlnAlaGlnAlaSerPheTrpThrAlaGluValAspLeuSerLysasp

20

QY

223

CTCCCTCACTGGAAACAGCTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCC

282

DB

21

LeuProHisTrpAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAla

40

QY

283

TTTTTTCGACCCAGTGGAAATGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAG

342

DB

41

PhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGlu

60

QY

343

GTCCAGGTTCCAGAGGCTCGCTTCTTCTATGGCTTTCAAAATTTCTCATCGAATGTTTAC

402

DB

61

ValGlnValProGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHis

80

QY

403

TCAGAGATGTACAGTTTGTGTAGACACTTACATCAGACATCCCAAGAAAAGGAAATTT

462

DB

81

SerGluMetTyrSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPhe

100

QY

463

TTATTTAATGCAATTCGAAACCATGCCCTATGTTTAAAGAAAAGCAGATTGGGCTTGGGA

522

DB

101

LeuPheAsnAlaIleGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArg

120

QY

523

TGGATAGCAGATAGAAAATCTACTTTTGGGAAAGAGTGTGGCTTTTGTCTCTGTAGAA

582

DB

121

TrpIleAlaAspArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGlu

140

QY

583

GGAGTTTTTCTTCTCAGGATCTTTTGTGCTATATTCTGGCTTAAAGAGAGAGGCTTATG

642

```

Db      141 GlyValPheSerGlySerPheAlaAlaPheTrpLeuLysLysArgGlyLeuMet 160
QY      643 CCAGGACTCACTTTTCCAAATGAACATCATCAGCAGAGATGAAGSACTTCACTGACTTT 702
Db      161 ProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPhe 180
QY      703 GCTTGCTGATGTTCCTCAATCTAGTAAATAAGCCTTCAGAGAAAGGGTCAGGAGATC 762
Db      181 AlaCysLeuMetPheGlnTyrLeuValAsnLysProGluGluArgValArgGluLeu 200
QY      763 ATTGTTGATGCTGCTCAAAATTGAGCAGGAGTCTTTTAAACAGAACCTTGCAGTGGCCTC 822
Db      201 IleValAspAlaValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeu 220
QY      823 ATTGGAATCAATTCATTTGATGAACAGTACATTCAGTTCCTGCTGACGATTCATT 882
Db      221 IleGlyMetAsnCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeu 240
QY      883 GTGGAATTCGGATTCTCAAGGTTTTCAGGCAGAAATCTTTTGTGATTTTATGGAAC 942
Db      241 ValGluLeuGlyPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsn 260
QY      943 ATTCTTTTGAAGAGAAACAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTT 1002
Db      261 IleSerLeuGluGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPhe 280
QY      1003 GCAGTTATGCCAGAAACACAGATACGCTTCACCTTGATGCAGATTTT 1053
Db      281 AlaValMetAlaGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 297

```

RESULT 5

```

Q75PY9 PRELIMINARY; PRT; 299 AA.
AC Q75PY9;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DE P53-inducible ribonucleotide reductase small subunit 2 long form.
GN Name=p53R2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RA Ugai H., Yokoyama K.K.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB163437; BAD11774.1; -.
DR GO; GO:0004748; P:ribonucleoside-diphosphate reductase activity; IEA.
DR GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonucleotide.
DR Pfam; PF00268; Ribonuc red sm; 1.
DR PROSITE; PS00368; RIBORED SMALL; 1.
SQ SEQUENCE 299 AA; 34528 MW; 4705C44389EB689B CRC64;

```

Alignment Scores:

```

Pred. No.:      3,03e-124      Length:      299
Score:          1514.00      Matches:      299
Percent Similarity: 85.19%      Conservative: 0
Best Local Similarity: 85.19%      Mismatches: 0
Query Match:      82.02%      Indels:      52
DB:              2          Gaps:      1

```

US-10-698-228-12 (1-1053) x Q75PY9 (1-299)

```

QY      1 ATGGGGGACCCCGAAGGCGGAGCGCGGGCTGGATCAGATGAGATCATCTTCA 60
Db      1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAsp----- 15
QY      61 GACACCAACGAAAGTGAATAAGTCAATGAAGACCACTCCTTAAGAAAGAGTTCCTCC 120

```

```

Db      15 ----- 15
QY      121 CGGTTTGTCACTTTTCCAATCCAGTACCCTCATATTTGAAATGTATATAACAGGCACAG 180
Db      15 ----- 15
QY      181 GCTTCTTCTTGACAGCAGAGAGAGGTTGACTTATCAAAAGGATCTCCCTCAGTGGAAACAG 240
Db      16 -----GluValAspLeuSerLysAspLeuProHisTrpAsnLys 28
QY      241 CTTAAACAGATGAGAAGTACTTCACTCTCATCTTACGCTTTTTCAGCCAGTGTAT 300
Db      29 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 48
QY      301 GGAATTCCTAATGAAATTTGGTCGAGCGCTTTAGTCAGGAGGTCACAGGTTCCAGAGGCT 360
Db      49 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 68
QY      361 CGCTGTTTCTATGCTTTTCAAAATCTCATCCAGAAATGTTCTACTCAGAGATGTACAGTTTG 420
Db      69 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 88
QY      421 CTGATACACACTTATACATCAGAGATCCCAAGAAAGGGAATTTTATTAAATGCAATGAA 480
Db      89 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 108
QY      481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCGCTTCCGATGGATGATGACAGATGAAAA 540
Db      109 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 128
QY      541 TCTACTTTTGGGAAAGAGTGGTGGCTTTCTGCTGTAGAGGAGTCTTCTTCTCAGCA 600
Db      129 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 148
QY      601 TCTTTTCTGCTGCTATATTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Db      149 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 168
QY      661 AATGAATCATCAGCAGATGAAGGACTTCACCTGTGACTTTTGTGCTGCTGATGTTCCAA 720
Db      169 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 188
QY      721 TACTTAGTAATAAGCTTCAGAGAAAGGCTCAGGAGATCATTTGATGCTGCTCATAA 780
Db      189 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 208
QY      781 ATTGAGCAGGAGTTTAAACAGAACCTTGCAGTTCGCCCTCATTGGAAATGAATGCATT 840
Db      209 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 228
QY      841 TTGATGAAACAGTACATGATGTTGTAGCTGACAGATPACTTGTGGAACCTTGGATCTCA 900
Db      229 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 248
QY      901 AAGGTTTTCAGGCAGAGAAATCTTTTGTATTTTATGAAAAATTTCTTTAGAGAGAAA 960
Db      249 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 268
QY      961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTCAGGTTTATGCAGAAACC 1020
Db      269 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 288
QY      1021 ACAGATAAACCTTCTTCACTTGGATGCAGATTTT 1053
Db      289 ThrAspAsnValPheThrLeuAspAlaAspPhe 299

```

RESULT 6

```

Q86YE3 PRELIMINARY; PRT; 285 AA.
AC Q86YE3;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)

```

DE Similar to ribonucleotide reductase M2 polypeptide (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC042468; AAH42468.1; -;
DR HSSP; P11157; 1XSN.
DR GO; GO:0004748; F:ribonucleoside-diphosphate reductase activity; IEA.
DR GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonucleotide reductase.
DR Pfam; PF00268; Ribonucleotide reductase; 1.
DR PROSITE; PS00368; RIBONUC_SMALL; 1.
FT NON TER 1
SQ SEQUENCE 285 AA; 33075 MW; C4E7D536479BC15B CRC64;

Alignment Scores:
Pred. No.: 9-87e-121 Length: 285
Score: 1474.00 Matches: 285
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 79.85% Indels: 0
DB: 2 Gaps: 0

US-10-698-228-12 (1-1053) x Q86YE3 (1-285)
QY 199 GAAGAGTTGACTTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAG 258
Db 1 GluGluValAspLeuSerGlySerLeuProHisTrpAsnLysLeuLysAlaAspGluLys 20

QY 259 TACTTCATCTCACAATCTAGCTTTTGGCAGCCAGTGGATGGAATTTGAATGAAAT 318
Db 21 TyrPheLeuSerHisLeuAlaPhePheAlaAlaSerAspGlyLeuValAsnGluAsn 40

QY 319 TTGGTGGAGCGCTTAGTACGAGGAGTGGAGTTCAGAGCTCGCTGTTCTTATGCGCTTT 378
Db 41 LeuValGluArgPheSerGlnGluValGlnValProGluAlaArgCysPheTyrGlyPhe 60

QY 379 CAAATCTCATCGAGATCTTCACTCAGAGATGATGAGTTTCTGCTGATGACACTTACATC 438
Db 61 GlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeuLeuIleAspThrTyrIle 80

QY 439 AGAGATCCCAAGAAAGGCAATTTTATTTAATGCAATTTGAAACCATGCGCTATGTTAAG 498
Db 81 ArgAspProLysArgGluPheLeuPheAsnAlaIleGluThrMetProTyrValLys 100

QY 499 AAAAAAGCAGATGGCGCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATG 558
Db 101 LysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLysSerThrPheGlyGluArg 120

QY 559 GTGGTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 618
Db 121 ValValAlaPheAlaAlaValGluGlyValPhePheSerGlySerPheAlaAlaIlePhe 140

QY 619 TGGCTTAAAGAGAGAGTCTTATGTCAGGAGTCACTTTTCCAAATGCACTCATCAGCAGA 678
Db 141 TrpLeuLysArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArg 160

QY 679 GATGAGGACTTCACTGTGACTTTGCTGCTGATGCTTCAATPACTTAGTAATAAGCCT 738
Db 161 AspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGlnTyrLeuValAsnLysPro 180

QY 739 TCAGAAAGAGGTCAGGAGATCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 798
Db 181 SerGluGluArgValArgGluIleIleValAspAlaValLysIleGluGlnGluPheLeu 200

QY 799 ACAGAGCGCTTGCAGTTCGCTTCAATGGAATGCAATTTGATGATGATGATGATGATGAT 858
Db 199 ACAGAGCGCTTGCAGTTCGCTTCAATGGAATGCAATTTGATGATGATGATGATGATGAT

Db 201 ThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIleLeuMetLysGlnTyrIle 220
QY 859 GAGTTTGTAGCTCAGACATTTCTGGAATCTTCTCAAGAGTTTCTCAGGSCAAA 918
Db 221 GluPheValAlaAspArgLeuValGluLeuGlyPheSerLysValPheGlnAlaGlu 240
QY 919 AATCCTTTTGTATTTATGGAACCATTTCTTTAGAGGAAAAACAAATTTCTTTGAGAAA 978
Db 241 AsnProPheAspPheMetGluAsnIleSerLeuGluGlyLysThrAsnPheGluLys 260
QY 979 CGAGTTTCAGATGATCAGCTTTTTCAGATGATGATGATGATGATGATGATGATGATGAT 1038
Db 261 ArgValSerGluTyrGlnArgPheAlaValMetAlaGluThrThrAspAsnValPheThr 280
QY 1039 TTGGATGTCAGATTTT 1053
Db 281 LeuAspAlaAspPhe 285

RESULT 7
RIR2 HUMAN STANDARD; PRT; 389 AA.
ID RIR2 HUMAN
AC P31350;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Ribonucleoside-diphosphate reductase M2 chain (EC 1.17.4.1)
DE (Ribonucleotide reductase small chain).
GN Name=RRM2; Synonym=RR2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast carcinoma;
RX MEDLINE=92329977; PubMed=1627826;
RA Pavloff N., Rivard D., Masson S., Shen S.-H., Mes-Masson A.-M.;
RT "Sequence analysis of the large and small subunits of human
ribonucleotide reductase."
RL DNA Seq. 2:227-234 (1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21974633; PubMed=11978970;
RA Zhou B., Yen Y.;
RT "Characterization of the human ribonucleotide reductase M2 subunit
gene; genomic structure and promoter analyses."
RL Cytogenet. Cell Genet. 95:52-59 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle, and Skin;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
RA Sapleton L., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Heiton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -1- FUNCTION: Provides the precursors necessary for DNA synthesis.
CC -1- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate +

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87089677; PubMed=3025593;
RA Thelander L., Berg P.;
RT "Isolation and characterization of expressible cDNA clones encoding
RL the M1 and M2 subunits of mouse ribonucleotide reductase.";
RM Mol. Cell. Biol. 6:3433-3442(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90060004; PubMed=2684652;
RA Thelander M., Thelander L.;
RT "Molecular cloning and expression of the functional gene encoding the
RL M2 subunit of mouse ribonucleotide reductase: a new dominant marker
RT gene.";
RL EMBL J. 8:2475-2479(1989).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 65-352.
RX MEDLINE=97030707; PubMed=8876648; DOI=10.1006/jmbi.1996.0546;
RA Kauppi B., Nielsen B.B., Ramaswamy S., Larsen I.K., Thelander M.,
RA Thelander L., Eklund H.;
RT "The three-dimensional structure of mammalian ribonucleotide reductase
RL protein R2 reveals a more-accessible iron-radical site than
RL Escherichia coli R2.";
RL J. Mol. Biol. 262:706-720(1996).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 65-352.
RX MEDLINE=22206532; PubMed=12087093; DOI=10.1074/jbc.M203358200;
RA Strand K.R., Karlsson S., Andersson K.K.;
RT "Cobalt substitution of mouse R2 ribonucleotide reductase as a model
RL for the reactive diferrrous state. Spectroscopic and structural
RL evidence for a ferromagnetically coupled dinuclear cobalt cluster.";
RL J. Biol. Chem. 277:34229-34238(2002).
RN [5]
RP STRUCTURE BY NMR OF 384-390.
RX MEDLINE=96069780; PubMed=7583667;
RA Fisher A.L., Laub P.B., Cooperman B.S.;
RT "NMR structure of an inhibitory R2 C-terminal peptide bound to mouse
RL ribonucleotide reductase R1 subunit.";
RL Nat. Struct. Biol. 2:951-955(1995).
CC -!- FUNCTION: Provides the precursors necessary for DNA synthesis.
CC -!- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate +
CC thioresoxin disulfide + H(2)O = ribonucleoside diphosphate +
CC thioresoxin.
CC -!- COFACTOR: Binds 2 iron ions per subunit.
CC -!- PATHWAY: DNA replication pathway; first step.
CC -!- SUBUNIT: Heterodimer of a large (M1) and a small chain (M2).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- MISCELLANEOUS: Two distinct regulatory sites have been defined:
CC the specificity site, which controls substrate specificity, and
CC the activity site which regulates overall catalytic activity. A
CC substrate-binding catalytic site, located on M1, is formed only in
CC the presence of the second subunit M2.
CC -!- SIMILARITY: Belongs to the ribonucleoside diphosphate reductase
CC small chain family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M14223; AAA40062.1; -;
CC EMBL; X15666; CAA33707.1; -;
CC PIR; S06735; S06735
CC PDB; 1AFT; NMR; @=383-390.
CC PDB; 1HON; X-ray; A=1-390.
CC PDB; 1HOO; X-ray; A=1-390.
CC PDB; 1XSM; X-ray; @=1-390.
CC MGD; MGI:98181; Rrm2.
CC GO; GO:0005515; P:protein binding; IPI.
DR GO; GO:0009262; P:deoxyribonucleotide metabolism; IDA.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonucl_redctse.
DR Pfam; PF00268; Ribonuc_red_sm; 1.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
KW 3D-structure; DNA replication; Iron; Metal-binding; Oxidoreductase.
RL METAL 139 139 Iron 1.
FT METAL 170 170 Iron 1 and 2.
FT METAL 173 173 Iron 1.
FT METAL 173 173 Iron 2.
FT METAL 233 233 Iron 1 and 2.
FT METAL 267 267 Iron 1 and 2.
FT METAL 270 270 Iron 2.
FT ACT_SITE 177 177 By similarity.
FT HELIX 68 70
FT TURN 72 74
FT HELIX 89 99
FT TURN 100 101
FT HELIX 105 107
FT TURN 112 113
FT HELIX 114 118
FT TURN 119 119
FT HELIX 122 148
FT TURN 149 149
FT HELIX 150 153
FT HELIX 157 184
FT HELIX 188 195
FT TURN 196 196
FT HELIX 197 200
FT HELIX 202 204
FT HELIX 205 216
FT TURN 218 219
FT HELIX 222 235
FT TURN 236 237
FT HELIX 238 249
FT TURN 250 251
FT HELIX 254 279
FT TURN 280 280
FT HELIX 287 306
FT TURN 307 308
FT HELIX 311 314
FT TURN 315 315
FT HELIX 318 335
FT TURN 336 337
FT HELIX 349 351
SQ SEQUENCE 390 AA; 45095 MW; AC7ACC4FAP8A4A2P CRC64;
Alignment Scores:
Pred. No.: 5,49e-118 Length: 390
Score: 1443.00 Matches: 269
Percent Similarity: 90.62% Conservative: 40
Best Local Similarity: 78.89% Mismatches: 26
Query Match: 78.17% Indels: 6
DB: 1 Gaps: 1
US-10-698-228-12 (1-1053) x RIR2_MOUSE (1-390)
QY 49 AGATCATCTTCACACACACAGCAAGTGAATAAAGTCA-----AAT 90
Db 50 ArgIlePheGlnAspSerAlaGluLeuGluSerLysAlaProThrAsnProSerValGlu 69
QY 91 GAAGACCACTCTTAAGAAAGATTCTCGCGGTTTGTTCATCTTTCCCAATCCAGTACCCT 150
Db 70 ASPLGLProuleuLeuArgGluAsnProArgPheValValPheProIleGluTyrHis 89
QY 151 GATATTGAAATGTATAAACAGGCAGCTTCTTCTGGACAGCAGAGAGGTTGAC 210
Db 90 ASPIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluValAsp 109
QY 211 TTATCAAGGATCTCCCTCACTGGAACAGCTTAAAGCAGATGAGAACTACTCATCTCT 270
Db 110 LeuSerLysAspIleGlnHisTrpGluAlaLeuLysProAspGluArgHisPheIleSer 129
QY 271 CACATCTTAGCCCTTTTTCGACGCCAGTGTGAATTGTAATGAAATTTGGTGGAGCGC 330

Db 211 IleSerAspLysGlnIaThrTyrGlyGluArgValValaPheAlaValaGluGly 230
 QY 586 GTTTCTCTCAGGATCTTTGCTGCTATATCTGGCTAAAGAGAGAGCTTATGCGCA 645
 Db 231 IlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArgGlyLeuMetPro 250
 QY 646 GGAATCCTTTTCCAATGAATCATCAGCAGAGATGAAGGACTTCACTGCTGCTTTGCT 705
 Db 251 GlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAla 270
 QY 706 TGCTGATGTTCCAACTACTAGTAATAAGCTTCAAGAAAGGTCAGGAGATCATTT 765
 Db 271 CysLeuMetPheArgHisLeuValHisLeYProSerGluGluArgValValGlnLeuIle 290
 QY 766 GTTGATGCTGTCAAAATTGAGCAGGAGTGTTTTAAACAGAGCCCTTGCAGTTGCCCTCAT 825
 Db 291 ThrAsnAlaValGlnIleGluGlnGluPheLeuThrGluAlaLeuProValAsnLeuIle 310
 QY 826 GGAATGAATTTGCAATTTGATGAACACAGTACATTCAGTTGTAGCTGACAGATTAATTG 885
 Db 311 GlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuLeu 330
 QY 886 GAACTTGGATCTCAAGGTTTTCAGCGCAGAAATCCTTTGATTTTATGAAAACATT 945
 Db 331 GluLeuGlyPheAsnLysValPheLysAlaSerAsnProPheAspPheMetGluAsnIle 350
 QY 946 TCTTTAGAGGAAACAAATTTCTTTGAGAAACGATTTTCAGATATCAGCGTTTGTGCA 1005
 Db 351 SerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGluTyrGlnLysMetGly 370
 QY 1006 GTTATGCGCAGAACACACAGATACGCTTTCACCTTGGATGTCAGATTTT 1053
 Db 371 ValMetSerLysProLysAspAsnThrPheThrLeuAspAlaAspPhe 386
 RESULT 11
 Q8AVY2 PRELIMINARY; PRT; 386 AA.
 ID Q8AVY2 AC Q8AVY2; DT 01-MAR-2003 (TREMBlrel. 23, Created) DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update) DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update) DE MGS2676 protein.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner K.H., Schenken C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalek U., Smaluk D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.

RC TISSUE=Embryo;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative."
 RL Dev. Dyn. 225:384-391(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S., Strausberg R.;
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC041209; AAH41209.1; -;
 DR HSP; P11157; IXS.
 DR GO; GO:0004748; P:ribonucleoside-diphosphate reductase activity; IEA.
 DR GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.
 DR InterPro; IPR009078; Ferritin/RR like.
 DR InterPro; IPR000358; Ribonuclease.
 DR Pfam; PF00268; Ribonuclease; 1.
 DR PROSITE; PS00368; RIBOED SMALL; 1.
 SQ SEQUENCE 386 AA; 44635 MW; 4635FCC9BB18EA32 CRC64;
 Alignment Scores:
 Pred. No.: 1,4e-116 Length: 386
 Score: 1427.00 Matches: 266
 Percent Similarity: 89.88% Conservative: 36
 Best Local Similarity: 79.17% Mismatches: 30
 Query Match: 77.30% Indels: 4
 DB: 2 Gaps: 1
 US-10-698-228-12 (1-1053) x Q8AVY2 (1-386)
 QY 58 TCAGACCAACCAAGTGAATAAGTCAAT-----GAAGAGCCTCTCTA 105
 Db 51 ThrGluTrpLysSerLysAlaProLysAsnProArgLeuGluAspGluProLeuLeu 70
 QY 106 AGAAGAGATTCTCGCGGTTTGTCTTCCATCCAGTACCTGATATTGGAAATG 165
 Db 71 LysAspAsnProHisArgPheValIlePheProLysGlnTyrHisAspIleTrpGlnMet 90
 QY 166 TATAACACGACACAGGCTTCTTCTGACAGCAGAGAGGTTTGACTTATCAAGGATCTC 225
 Db 91 TyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeu 110
 QY 226 CTTACTGGAACCAAGCTTAAAGCAGATGAGAAGTCTCTCATCTCTCATCTTACCTTT 285
 Db 111 GlnHisTrpGluSerLeuLysLysGluGluLysTyrPheIleSerHisValLeuAlaPhe 130
 QY 286 TTTGACACGACGATGGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGCTG 345
 Db 131 PheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerLysGluVal 150
 QY 346 CAGGTTCCAGACGCTCGCTTCTTCTATGGCTTTTCAAAATTTCTCATCGAATGTTCACTCA 405
 Db 151 GlnValThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMetGluAsnIleHisSer 170
 QY 406 GAGATGTACAGTTGCTGTATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTT 465
 Db 171 GluMetTyrSerLeuLeuIleAspThrTyrValLysAspProLysGluArgGluTyrLeu 190
 QY 466 TTTAATGCAATTCAGAACCATGCGCTTATGTAAGAAAAGCAGATTTGGCTTGGATGG 525
 Db 191 PheAsnAlaIleGluThrLeuProCysValLysLysLysAlaAspTrpAlaLeuHisTrp 210
 QY 526 ATAGCAGATGAAATCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTGTAGAAGCA 585
 Db 211 IleGlyAspLysGlnAlaThrPheGlyGluArgValValAlaPheAlaValaGluGly 230
 QY 586 GTTTCTCTCAGGATCTTTTGTGCTGCTATATCTGGCTTAAAGAGAGAGGCTTTATGCA 645
 Db 231 IlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArgGlyLeuMetPro 250
 QY 646 GGAATCCTTTTCCAATGAATCATCAGCAGAGATGAAGGACTTCACTGCTGCTTTGCT

Db 251 GlyLeuThrPheSerAsnGluLeuLeuSerArgSpGluGlyLeuHisCysAspPheAla 270
QY 706 TGCCTGATGTCCTCAATAGCTTAATAAGCTTTCAGAGAAAGGTCAGGAGATCAT 765
Db 271 CysLeuMetPheLysLeuLeuLeuArgLysProSerGluGluValGluLeuLeu 290
QY 766 GTTGATGCTGCTCAAAATTCAGACGAGGAGTTTAAACAGAGACCTTCGACGTCCTCAT 825
Db 291 ThrAspAlaValGlnIleGluGlnGluPheLeuThrGluAlaLeuProValAsnLeuLeu 310
QY 826 GGAATGAATTCATTTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTCATTCGTG 885
Db 311 GlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuLeu 330
QY 886 GAATTCGATTCCTCAAGCTTTTTCAGGACGAAATCCTTTTGAATTTATGGAACATTC 945
Db 331 GluLeuGlyPheLysLysValPheLysAlaThrAsnProPheAspPheMetGluAsnIle 350
QY 946 TCTTTAGAGGAGAAACAAATTCCTTTGAGAACGAGTTTCAGAGTATCAGCGTTTGTGCA 1005
Db 351 SerLeuGluGlyLysThrAsnPhePheGluLysLysValGlyGluTyrGlnLysMetGly 370
QY 1006 GTTATGGCAGAACCCACAGATAACGCTTCACCTTCGATGCAGATTTT 1053
Db 371 ValMetSerLysAlaLysAspAsnThrPheThrLeuAspAlaAspPhe 386
RESULT 12
RIR2_BRARE STANDARD; PRT; 386 AA.
AC P79733;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 25-OCT-2004 (Rel. 45, Last annotation update)
DE Ribonucleoside-diphosphate reductase M2 chain (EC 1.17.4.1)
DE (Ribonucleotide reductase protein R2 class I).
GN Name=rnm2; Synonyms=r2;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97137859; PubMed=8983196;
RA Mathews C.Z., Sjoberg B.-M., Karlsson M.;
RT "Cloning and sequencing of cDNAs encoding ribonucleotide reductase
RT from zebrafish Danio rerio."
RL Mol. Mar. Biol. Biotechnol. 5:284-287(1996).
CC -!- FUNCTION: Provides the precursors necessary for DNA synthesis.
CC -!- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate +
CC thioresoxin disulfide + H(2)O = ribonucleoside diphosphate +
CC thioresoxin.
CC -!- COFACTOR: Binds 2 iron ions per subunit (By similarity).
CC -!- PATHWAY: DNA replication pathway; first step.
CC -!- SUBUNIT: Heterodimer of a large (M1) and a small chain (M2).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the ribonucleoside diphosphate reductase
CC small chain family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U57965; AAB37103.1; -.
DR HSSP; P11157; 1XSM.
DR ZFIN; ZDB-GENE-990415-25; rnm2.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonucl_reductase.

DR Pfam; PF00368; Ribonuc_red_sm; 1.
KW PROSITE; PS00368; RIBRED_SMALL; 1.
KW DNA replication; Iron; Metal-binding; Oxidoreductase.
FT METAL 135 135 Iron 1 (By similarity).
FT METAL 166 166 Iron 1 and 2 (By similarity).
FT METAL 169 169 Iron 1 (By similarity).
FT METAL 229 229 Iron 2 (By similarity).
FT METAL 263 263 Iron 2 (By similarity).
FT METAL 266 266 Iron 2 (By similarity).
FT ACT_SITE 173 173 By similarity.
SQ SEQUENCE 386 AA; 44594 MW; C16846FB57F9F4E CRC64;
Alignment Scores:
Pred. No.: 1-89e-116 Length: 386
Score: 1425-50 Matches: 268
Percent Similarity: 89-41% Conservative: 36
Best Local Similarity: 78-82% Mismatches: 31
Query Match: 77-22% Indels: 5
DB: 1 Gaps: 1
US-10-698-228-12 (1-1053) x RIR2_BRARE (1-386)
QY 49 AGATCATCTTCAGACACCAACGAAAGTGAATAAAGTCA-----AATGAA 93
Db 47 ArgLysIlePheAspGluSerGluGlyGlnSerLysAlaLysLysGlyAlaValGluGlu 66
QY 94 GAGCCACTCTAAGAAAGAGATTCTCGCCGGTGTTCATCTTCCAAATCCAGTACCCGTAT 153
Db 67 GluProLeuLeuLysGluAsnProHisArgPheValIlePheProIleGlnTyrHisAsp 86
QY 154 ATTTGAAAATGTATAAACAGGACACAGGCTTCCTTCGGACAGCAGAGAGGTTTGACTTA 213
Db 87 IleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAspLeu 106
QY 214 TCAAGAGATCTCCCTCACTCGAAACAGCTTAAAGACAGATGAGAGTACTTCATCTCAC 273
Db 107 SerLysAspLeuGlnHisTrpAspSerLeuLysAspGluArgTyrPheIleSerHis 126
QY 274 ATCTTAGCCTTTTTCAGCCAGCTGAGTAATGTAATGTAATGTAATGTAATGTAATGTA 333
Db 127 ValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArgPhe 146
QY 334 AGTCAGGAGGTGCGAGTTCAGAGGCTCGCTGTTCTTATGCTTCAAAATTCATCCAG 393
Db 147 ThrGlnGluValGlnValThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMetGlu 166
QY 394 AATGTTCACTCAGAGATGTACAGTTTGTCTGATGACACTTACATCAGAGATCCCAAGAA 453
Db 167 AsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspSerLysGlu 186
QY 454 AGGGAATTTTATTAATGCAATTAACCATGCCCTATGTTAAGAAAGACAGATGG 513
Db 187 ArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAspTrp 206
QY 514 GCCTTCGGATGGATAGCAGATGAGAAATCTACTTTTGGGAAAGAGTGGCTTGTGCT 573
Db 207 AlaLeuAsnTrpIleGlyAspLysAsnAlaArgTyrGlyGluArgValAlaPheAla 226
QY 574 GCTGTAGAAGGAGTTTCTTCTCAGGATCTTTGCTGCTATATTCGCTAAAGAGAGA 633
Db 227 AlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysArg 246
QY 634 GGTCTTATGCCAGACTCAGCTTTTTCATGAATCACTCATCAGCAGATGAAGGACTTCAC 693
Db 247 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHis 266
QY 694 TGTGACTTGTGCTGCTGATGTTCCCAATCTACTAGTAATAAGCCTTCAGAGAAAGGCTC 753
Db 267 CysAspPheAlaCysLeuMetPheLysHisLeuIleAsnLysProSerGluGluThrVal 286
QY 754 AGGAGATCATTTGTTGATGCTGTCAAAATTCAGCAGGAGTGTTTTAAACAGAGCTTGCCA 813
Db 287 LysLysIleIleMetAsnAlaValGluIleGluGlnGluPheLeuThrAspAlaLeuPro 306

Qy	898	TCAAAGTCTTTTCAGGCAGAAAAATCCCTTTTGATTTATGAAAAACATTTCTTTAGAAAGGA	957
		::: ::: :::	
Db	339	AsnLysIlePheLysValGluAsnProPheAspPheMetGluAsnIleSerLeuGluGly	358
Qy	958	AAACAAATCTCTTGTGAACACGAGTTTCACAGTATCAGCGTTTTCAGTATGGCAGAA	1017
Db	359	LysThrAsnPhePheGluLysArgValGlyGlyTyrGlnArgMetGlyValMetSer---	377
Qy	1018	ACCACAGATAACGCTCTTCACCTTCGATGCAGATTTT	1053
Db	378	-----AsnSerPheThrLeuAspAlaAspPhe	386
RESULT 14			
Q6DI44		PRELIMINARY;	PRT; 386 AA.
ID	Q6DI44		
AC	Q6DI44		
DT	25-OCT-2004	(TrEMBLrel. 28, Created)	
DT	25-OCT-2004	(TrEMBLrel. 28, Last sequence update)	
DT	25-OCT-2004	(TrEMBLrel. 28, Last annotation update)	
DE	Ribonucleotide reductase M2 polypeptide.		
DE	Name=rm2;		
OS	Brachydanio rerio (Zebrafish) (Danio rerio).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;		
OC	Cyprinidae; Danio.		
OX	NCBI_TaxID=7955;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RS	TISSUE=Embryo;		
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;		
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,		
RA	Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,		
RA	Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,		
RA	Jones S.J., Marra M.A.;		
RT	"Generation and initial analysis of more than 15,000 full-length human		
RT	and mouse cDNA sequences.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RS	TISSUE=Embryo;		
RC	TISSUE=Embryo;		
RL	Strausberg R.;		
RL	Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC075746; AAH75746.1; -		
DR	GO; GO:0004748; F:ribonucleoside-diphosphate reductase activity; IEA.		
DR	GO; GO:0009146; P:deoxyribonucleoside diphosphate metabolism; IEA.		
DR	InterPro; IPR009078; Ferritin/RR like.		
DR	InterPro; IPR000358; Ribonuc_redctase.		
DR	Pfam; PF00268; Ribonuc_red_sm; 1.		
DR	PROSITE; PS00369; RIBORED_SMALL; 1.		
SQ	SEQUENCE 386 AA; 44567 MW; 2CCACD442ADD1566 CRC64;		

US-10-598-228--12 (1-1053) x QGD144 (1-386)

Qy	49	AGATCATCTTTACAGACACCACGAAGAAGTGAATAAAGTCAC-----AATGAA	93
Db	47	ArgLysIlePheAspGluSerGluGlyGlnSerLysAlaLysLysGlyAlaValGluGlu	66
Qy	94	GAGCCACTCCTAGAAGAAGTCTCCCGGTGGTGTCATCTTCCCAATCCAGTACCTGAT	153
Db	67	GluProLeuLeuLysGluAsnProHisArgPheValIlePheProIleGlnTyriHisAsp	86
Qy	154	ATTTTGAAAATGATATAAACAGGCACGAGGCTTCCTCTGGACACGACAAGAGGTGTGACTTA	213
Db	87	IleTrpGlnMetIryrLysLysAlaGluAlaSerPheTrpThrIaGluGluValaAspLeu	106
Qy	214	TCAAAGSAGTCTCCCTCACATGGAAACAAGCTTTAAAGCAGATCAGAAAGTACTTCATCTCTCAC	273
Db	107	SerLysAspLeuGlnHisTrpAspSerLeuLysAspGluGluArgIyrPheIleSerHis	126
Qy	274	ATCTTAGCTTTTTTCGAGCCAGTGTGAATTTGTAATGAANAATTTGGTGGAGCGCTTTT	333
Db	127	ValLeuAlaPhePheAlaIaSerAspGlyIleValAsnGluAsnLeuValGluIarGphe	146
Qy	334	AGTCAGSAGTGCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTTCTCATCTCGAG	393
Db	147	ThrGlnGluValGlnValThrGluAlaArgCySPheTyGlyPheGlnIleAlaMetGlu	166
Qy	394	AATGTTTCACATCAGAGATGTACAGTTTGTCTGATAGACATTCATACAGAGATCCCAAGAAA	453
Db	167	AsnIleHisSerGluMetIyrSerLeuLeuIleAspThrIryrileLysAspSerLysGlu	186
Qy	454	AGGGAATTTTATTAATGCAATTTGAAAAACATGCCCTATGTGTTAAGAAAAAAGCAGATTGG	513
Db	187	ArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAspTrp	206
Qy	514	GCCTTCGATGGATACGACATGACAAAATCTACTTTTGGGGAAGAGTGGTGGCTTTGCT	573
Db	207	AlaLeuAsnTrpIleGlyAspLysAsnAlaArGtyrGlyGluArgValValAlaPheAla	226
Qy	574	GCTGTGAAGAGTGTTCCTCTCAGGATCTTTTGTGCTGCTATATTCYGGCTAAAGAGAAGA	633
Db	227	AlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArg	246
Qy	634	GGTCTTANGCCAGACTCCTTTTCCAATGAACCTCATCAGCAGAGATGAAGACTTTCAC	693
Db	247	GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgaspGluGlyLeuHis	266
Qy	694	TGTGACTTTGCTTGCCTGATGTTTCCAATCTAGTAAATAAGCTTTCAGAAGAAGGGTC	753
Db	267	CysAspPheAlaCysLeuMetPheLysHisLeuIleSerLysProserGluGluThrVal	286
Qy	754	AGGAGGATCATTTGATGCTGTCACAAATTTGACGAGGAGTTTTTAACAGAAGCTTTGCCA	813
Db	287	LysLysIleIleMetAsnAlaValGluIleGluGlnGluPheLeuThrAspAlaLeuPro	306
Qy	814	GTTGGCCTCATTTGGAATGAATTTGATGAACACAGTACATTCAGTTTGTAGCTGAC	873
Db	307	ValLysLeuIleGlyMetAsnCysAspLeuMetLysGlnTrIyleGluPheValAlaAsp	326
Qy	874	AGATTACTTGTGAACCTTGATTTCTCAAAGTTTTTCAGCAGAAAAATCCTTTTGATTTT	933
Db	327	ArgLeuLeuLeuGluLeuGlyPheAspLysValTyArgValGluAsnProPheAspPhe	346
Qy	934	ATGGAAAACATTTCTTTAGAAGCAAAAACAAATTTCTTTGAGAAAACGAGTTTCAGAGTAT	993
Db	347	MetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGluTy	366
Qy	994	CAGCGTTTTGCAGTTATGGCAGAAAACACAGATAACGCTCTTCACCTTGGATGCAGATTTT	1053
Db	367	GlnArgMetGlyValMetSerGlyThrThrAspAsnThrPheThrLeuAspAlaAspPhe	386

RESULT 15
Q6P876

PRELIMINARY; PRT; 378 AA.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 30, 2005, 06:01:36 ; Search time 15.5 Seconds
(without alignments)
10142.641 Million cell updates/sec

Title: US-10-698-228-12

Perfect score: 1846

Sequence: 1 atggcgaccggaaagcc.....tcacctggatgcagatttt 1053

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10698228/runat 26102005.100609 4010/app_query.fasta 1.2446
-DB=Issued Patents AA -QFMT=fastan -SUFFIX=n2p.ra -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NOR=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228 @CGN 1 1 32 .runat 26102005.100609 4010 -NCPU=6 -ICPU=3
-NO MNAP -LARGEQUERY -NEG SCORES=0 -WAIT -DISPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	1821	98.6	351	4	US-10-019-733-1	Sequence 1, Appli
2	1454	78.8	389	4	US-10-019-733-5	Sequence 5, Appli
3	1454	78.8	389	4	US-09-949-016-6016	Sequence 6016, Ap
4	1454	78.8	453	4	US-09-949-016-7896	Sequence 7896, Ap
5	1131	61.1	411	4	US-09-248-796A-18017	Sequence 18017, A
6	1072.5	58.1	274	1	US-08-307-499-19	Sequence 19, Appl
7	1072.5	58.1	274	3	US-09-299-268-19	Sequence 19, Appl
8	974.5	52.8	322	1	US-08-136-743B-2	Sequence 2, Appli
9	969.5	52.5	360	4	US-09-248-796A-18018	Sequence 18018, A
10	768	41.6	147	3	US-08-905-223-358	Sequence 358, App
11	599	32.4	149	4	US-09-270-767-46236	Sequence 46236, A
12	495	26.8	323	4	US-09-902-540-10006	Sequence 10006, A

13	381.5	20.7	445	4	US-09-328-352-4714	Sequence 4714, Ap
14	377.5	20.4	476	4	US-09-352-991A-19358	Sequence 19358, A
15	335	18.1	365	4	US-09-198-452A-1059	Sequence 1059, Ap
16	335	18.1	380	4	US-09-438-185A-986	Sequence 986, App
c 17	299.5	16.3	121	1	US-08-307-499-7	Sequence 7, Appli
c 18	299.5	16.3	121	3	US-09-299-268-7	Sequence 7, Appli
19	298	16.1	314	3	US-09-080-044-3	Sequence 3, Appli
20	298	16.1	314	3	US-09-531-857A-3	Sequence 3, Appli
21	294	15.9	56	3	US-08-905-223-362	Sequence 362, App
22	220	11.9	382	4	US-09-543-681A-4239	Sequence 4239, Ap
23	203.5	11.0	412	4	US-09-489-039A-8180	Sequence 8180, Ap
24	201.5	10.9	401	4	US-09-540-236-3566	Sequence 3566, Ap
25	200	10.8	325	4	US-09-107-532A-4443	Sequence 4443, Ap
26	195.5	10.6	320	3	US-09-092-437-2	Sequence 2, Appli
27	195.5	10.6	320	4	US-09-583-110-3858	Sequence 3858, Ap
28	195.5	10.6	320	4	US-09-107-433-3136	Sequence 3136, Ap
29	195	10.6	339	4	US-09-543-681A-5069	Sequence 5069, Ap
30	169.5	9.2	323	4	US-09-042-775-2	Sequence 2, Appli
31	167	9.0	343	3	US-09-134-001C-5484	Sequence 5484, Ap
32	163	8.8	334	4	US-09-134-000C-5756	Sequence 5756, Ap
33	161	8.7	311	4	US-09-710-279-1260	Sequence 1260, Ap
34	157	8.5	324	2	US-08-813-940-2	Sequence 2, Appli
35	156	8.5	391	4	US-09-489-039A-10121	Sequence 10121, A
36	153	8.3	104	4	US-09-230-485-5	Sequence 5, Appli
37	138	7.5	321	2	US-08-813-940-6	Sequence 6, Appli
38	127	6.9	52	3	US-08-917-299-21	Sequence 21, Appl
39	127	6.9	52	3	US-09-422-662-21	Sequence 21, Appl
40	127	6.9	52	4	US-09-730-763-21	Sequence 21, Appl
41	127	6.9	52	4	US-09-429-370-21	Sequence 21, Appl
42	103	5.6	555	3	US-09-173-151A-32	Sequence 32, Appl
43	100.5	5.4	382	4	US-09-134-000C-5889	Sequence 5889, Ap
44	99	5.4	233	3	US-08-993-825-2	Sequence 2, Appli
45	98.5	5.3	896	4	US-09-270-767-46130	Sequence 46130, A

ALIGNMENTS

RESULT 1
US-10-019-733-1
; Sequence 1, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 1
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-1

Alignment Scores:			
Pred. No.:	4,29e-210	Length:	351
Score:	1821.00	Matches:	351
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	98.65%	Indels:	0
DB:	4	Gaps:	0

US-10-698-228-12 (1-1053) x US-10-019-733-1 (1-351)

QY 1 ATGGGCGACCCGGAAAGCGGCGCGGCTGGATCAGATGAGATCATCTTCA 60

```
Db      1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY      61 GACACCAAGCAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAGAGTTCTCCG 120
Db      21 AspThrAsnGluSerGluIleYsSerAsnGluGluProLeuLeuArgLysSerArg 40
QY      121 CGTTTGTGCATCTTCCCAATCCAGTACCCTGATATTGTGAAAATGTATAAACAAGGCACAG 180
Db      41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysGlnAlaGln 60
QY      181 GCTTCCTCTGACACAGCAAGAGTTGACTTATCAAGAGTCTCCCTCACTGGACAAG 240
Db      61 AlaSerPheThrTrpAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY      241 CTTAAGCAGATGAGAGTACTTCTCTCTCATCTTACATCTTACCTTTTTCAGCCAGTAT 300
Db      81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY      301 GGAATTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCAGAGGCT 360
Db      101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY      361 CGCTGTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db      121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY      421 CTGATAGACACTTACATCAGACATCCCAAGAAAGGAATTTTATTAATGCAATTGAA 480
Db      141 LeuIleAspThrTyrIleArgAspProLysArgGluPheLeuPheAsnAlaIleGlu 160
QY      481 ACCATGCCCTATGTTAAGAAAAAGCAGATTCGGCTTCGGATGGATAGCAGATGAAAA 540
Db      161 ThrMetProTyrValLysLysLysAlaAspThrAlaLeuArgTrpIleAlaAspArgLys 180
QY      541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGAGAGTTTCTTCTCAGCA 600
Db      181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY      601 TCTTTTGTGCTATATTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
Db      201 SerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY      661 AATGAACCTATCAGCAGAGATCAAGACTTCACTGTGACTTGTCTGCTGCTGATGTTCCAA 720
Db      221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY      721 TACTTAGTAAATTAAGCCTTTCAGAGAAAGGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db      241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260
QY      781 ATTGACGAGGAGTTTAAACAGAGCCTTGGCAGTTGGCTCATTTGGAATGAATGCAAT 840
Db      261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY      841 TTGATGAACAGTACATTCAGTTTCTGCTGACAGATTTACTTGTGAACTTGGATCTCA 900
Db      281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY      901 AAGGTTTTTCAGGCAGAGAAAATCTTTGATTTTATGGAAAACATTTCTTTAGAGGAAAA 960
Db      301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY      961 ACAAAATTTCTTGAAACAGAGTTTTCAGAGTATCAGCGTTTTCAGAGTTTATGCGAGAAC 1020
Db      321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY      1021 ACAGATACGCTCTTACCTTGGATCGAGATTTT 1053
Db      341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 2

US-10-019-733-5

```
; Sequence 5, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 5
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
; US-10-019-733-5
```

Alignment Scores:

Pred. No.:	7,14e-166	Length:	389
Score:	1454.00	Matches:	269
Percent Similarity:	94.08%	Conservative:	33
Best Local Similarity:	83.80%	Mismatches:	19
Query Match:	78.76%	Indels:	0
DB:	4	Gaps:	0

US-10-698-228-12 (1-1053) x US-10-019-733-5 (1-389)

```
QY      91 GAAGAGCCACTCTCTAAGAAAGATTCTCGCGGTTTGTCTATCTTTCCAAATCCAGTACCCT 150
Db      69 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 88
QY      151 GATATTTGGAATAATGATAAAGCAGGACAGCTTCTCTGACAGCAGAGAGAGTTGAC 210
Db      89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
QY      211 TTATCAAAAGGATCTCCTCACTCGAAACAAGCTTAAAGCAGATGAGAAAGTACTTCTCTCT 270
Db      109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY      271 CACATCTTAGCCTTTTTCAGCCAGTATGGAATTTGAAATTTGAAATTTGGTGGAGCGC 330
Db      129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY      331 TTATGTCAGAGGTGCGAGTTCCAGAGGCTCGCTGTTTCTATGCTTTTCAAAATTTCTCATC 390
Db      149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY      391 GAGAAATGTTCACTCAGAGATGTACAGTTGCTGATAGACACTTACATCAGAGATCCCAAG 450
Db      169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY      451 AAAAGCGAAATTTTATTAAATGCAATTGAAACCATGCCCTATGTTTAAAGAAAAAGCAGAT 510
Db      189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
QY      511 TGGGCTTCGGATGGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCCTTT 570
Db      209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 228
QY      571 GCTGCTGTAGAAGAGTCTTCTCTCAGAGTCTTTTGTCTGCTATATTTCTGCTAAAGAG 630
Db      229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheThrPheLysLys 248
QY      631 AGAGGCTTTATGCGCAGGACTCACTTTTTCAAATGAACCTCATCAGCAGAGATGAAGACTT 690
Db      249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
```

```
QY 691 CACTGTGACTTTGCTTGGCTGATGTTCCAAATCTTAGTAATAAGCCTTCAGAAAGG 750
Db 269 HisCysAspPheAlaCysLeuMetPheLeuValHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGAGATCATTTGATGCTGTCAAAATTTGACAGGAGTTTAAACAGAGCCTTG 810
Db 289 ValArgGluLeuLeuLeuLeuAlaValArgGluGluGlnGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCTCATTTGGAATGAAATGCAATTTGATGAAACAGATPACATTTGAGTTTGTAGCT 870
Db 309 ProValLeuLeuLeuLeuLeuMetLysGlnThrLeuMetLysGlnThrLeuPheValAla 328
QY 871 GACAGATTACTTGTGGAATCTGGAATCTCAAGGTTTTCAGCAGAGAAATCCTTTTGAT 930
Db 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY 931 TTTATGGAACATTTCTTTAGAGAAACAAATTTCTTTGAGAAACGAGTTTCAGAG 990
Db 349 PheMetGluAsnLeuSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY 991 TATCAGCGTTTTCAGTTATGCGTATGCGTATGCGTATGCGTATGCGTATGCGTATGCGTAT 1050
Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY 1051 TTT 1053
Db 389 Phe 389

RESULT 3
US-09-949-016-6016
; Sequence 6016, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6016
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6016

Alignment Scores:
Pred. No.: 7,14e-166 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 4 Gaps: 0

US-10-698-228-12 (1-1053) x US-09-949-016-6016 (1-389)
QY 91 GAAGAGCCACTCTCAAGAAAGAGTTCTCGCGGTTGTCTATCTTCCAAATCAGTACCT 150
Db 69 AspGluProLeuLeuArgGluAsnProArgArgPheValIlePheProIleGluTyrHis 88
QY 151 GATATTGAAATGTATAAACAGGCACAGGCTTCTTCTGACACAGCAAGAGTTGAC 210
Db 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheThrPheAlaGluGluValAsp 108
QY 211 TTATCAAGAGTCTCCCTCACTGGACACAGCTTAAAGCAGATGAGAGTACTCTCT 270
Db 211 TTATCAAGAGTCTCCCTCACTGGACACAGCTTAAAGCAGATGAGAGTACTCTCT
```

```
Db 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTTACCTTTTTCAGCCAGCAGTATGGAATTTGAAATGAAATTTTGGTGGAGCC 330
Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY 331 TTTAGTCAGGAGGTGTCAGGCTCCAGGCTCGCTGTTTCTATGGCTTTCAAATTTCTCATC 390
Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGATGTTCTCCTCAGAGATGTACAGTTGCTGTGATAGACACTTACATCAGAGATCCCAG 450
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAAGGGAATTTTATTAATGCAATTAATGCAATTAATGCAATTAATGCAATTAATGCAATTA 510
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
QY 511 TGGCCCTTGGCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 570
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValValAlaPhe 228
QY 571 GCTGCTGTAGAGAGTTTCTTCTCAGGATCTTTTCTGCTGCTATATCTGCTGCTGCTGCTG 630
Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGTCTTATGCGCAGCAGTCTCCTTTTCCAAATTAAGTATGATGATGATGATGATGATGAT 690
Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY 691 CACTGTGACTTTGCTTGGCTGATGTTCCAAATCTTAGTAATAAGCCTTCAGAAAGG 750
Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGAGATCATTTGATGCTGTCAAAATTTGAGAGGAGTTTTCAGCAGAGCCTTG 810
Db 289 ValArgGluLeuLeuLeuLeuAlaValArgGluGluGlnGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCTCATTTGGAATGAAATGCAATTTGATGAAACAGATPACATTTGAGTTTGTAGCT 870
Db 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
QY 871 GACAGATTACTTGTGGAATCTGGAATCTCAAGGTTTTCAGCAGAGAAATCCTTTTGAT 930
Db 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY 931 TTTATGGAACATTTCTTTAGAGAAACAAATTTCTTTGAGAAACGAGTTTCAGAG 990
Db 349 PheMetGluAsnLeuSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY 991 TATCAGCGTTTTCAGTTATGCGTATGCGTATGCGTATGCGTATGCGTATGCGTATGCGTAT 1050
Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY 1051 TTT 1053
Db 389 Phe 389

RESULT 4
US-09-949-016-7896
; Sequence 7896, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
```


Db 201 CysAspPheAlaCysLeuMetPheLysHisLeuLeuHisProSerLysGluValIle 220
QY 754 AGGGAGATCATTTGATGCTCTCAAAATTTGACGAGGAGTTTAAACAGAGCCTTGCCA 813
Db 221 ThrSerIleIleIleAspAlaValAsnIleGluLysGluPheLeuThrValAlaIlePro 240
QY 814 GTTGCGCTCATTTGGAATGAATTTGATGAAACAGTACATTCAGTTTGTAGCTGAC 873
Db 241 ValAspLeuIleGlyMetAsnCysCysLeuMetSerGlnTyrIleGluPheValAlaAsp 260
QY 874 AGATTACTTGTGGAATTCGATTTGATGAAACAGTACATTCAGTTTGTAGCTGAC 873
Db 261 ArgLeuLeuThrGluLeuGlyCysGluLys 270

RESULT 7

US-09-299-268-19
; Sequence 19, Application US/09299268
; Patent No. 6217882
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Viuela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,268
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/901,127
; FILING DATE:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 274 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-299-268-19

Alignment Scores:
Pred. No.: 5,21e-120 Length: 274
Score: 1072.50 Matches: 204

Percent Similarity: 85.56% Conservative: 27
Best Local Similarity: 75.56% Mismatches: 38
Query Match: 58.10% Indels: 1
DB: 3 Gaps: 1
US-10-698-228-12 (1-1053) x US-09-299-268-19 (1-274)
QY 94 GAGCCACTCCTAAAGAAAGAGTCTCGCCGGTTTGTCTCATCTTTTCCAAATCCAGTACCCTGAT 153
Db 2 GluProIleLeuGlnGluSerAspSerArgPheValIlePheProIleLysTyrHisAsp 21
QY 154 ATTTGAAAATGTATAACAGCAGCAGGCTTCTGGACAGCAGCAGAGAGGTTGACTTA 213
Db 22 IleTrpLysMetTyrLysGlnSerValAlaSerPheTrpThrValGluGluValAspLeu 41
QY 214 TCAAGATCTCCCTCAGTGGAAACAGCTTAAGCAGATGAGAAGTACTTCTCTCAC 273
Db 42 SerLysAspLeuAspAspTrpAspLysLeuThrLysAspGluLysTyrPheIleLysHis 61
QY 274 ATCTTAGCCCTTTTTCAGCAGCTGATGAATTCCTAAATGAAATTTGCTGAGCGCTTT 333
Db 62 IleLeuAlaPhePheAlaSerSerAspGlyLeValAsnGluAsnLeuAlaGluArgPhe 81
QY 334 AGTCAGGAGGTGCAGGTTCCAGAGGCTCGCTGTTTCTATGCTTTTCAAAATTTCTCATCGAG 393
Db 82 TyrValAspValGlnCysSerGluAlaArgCysPheTyrGlyPheGlnIleAlaMetGlu 101
QY 394 AATGTTCACTCAGAGATGTACAGTTGCTGATAGACACTTACATCAGAGATCCCAAGAAA 453
Db 102 AsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrValArgAspAsnIleGlu 121
QY 454 AGGGAATTTTATTATGCAATTGAAACCATGCCCTATGTTAAAGAAAAAGCAGATTGG 513
Db 122 LysMetHisLeuPheAsnAlaIleGluThrMetGluCysValLysLysLysLysLysLys 141
QY 514 GCCTTCCGATGTAGCAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCCTTTGCT 573
Db 142 AlaArgLysTrpIleSerSerAsnLys--ValTyrGlyGluArgValValAlaPheAla 160
QY 574 GCTGTAGAGGAGTTTCTTCTCAGGATCTTTGCTGCTATATTTCTGCTGCTAAAGAGAGA 633
Db 161 AlaValGluGlyIlePhePheSerGlySerPheAlaAlaIlePheTrpIleLysLysArg 180
QY 634 GGTCTTATGCCAGGACTCATCTTTTCCAATGAATCTCATCAGCAGATGAGAGGACTTCAC 693
Db 181 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLysHis 200
QY 694 TGTGACTTTTCTGCTGCTGATGTTCCAAATCTTCAATTAAGCCTTCAGAAAAGGGGTC 753
Db 201 CysAspPheAlaCysLeuMetPheLysHisLeuHisProProSerLysGluValIle 220
QY 754 AGGGAGATCATTTGATGCTGTCAAAATTTGACGAGGAGTTTAAACAGAGCCTTGCCA 813
Db 221 ThrSerIleIleIleAspAlaValAsnIleGluLysGluPheLeuThrValAlaIlePro 240
QY 814 GTTGCGCTCATTTGGAATGAATTTGATGAAACAGTACATTCAGTTTGTAGCTGAC 873
Db 241 ValAspLeuIleGlyMetAsnCysCysLeuMetSerGlnTyrIleGluPheValAlaAsp 260
QY 874 AGATTACTTGTGGAATTCGATTTGATGAAACAGTACATTCAGTTTGTAGCTGAC 873
Db 261 ArgLeuLeuThrGluLeuGlyCysGluLys 270

RESULT 8

US-08-136-743B-2
; Sequence 2, Application US/08136743B
; Patent No. 5459063
; GENERAL INFORMATION:
; APPLICANT: Barry S. Cooperman, Harvey Rubin,
; APPLICANT: Jerome Salem, and Alison L. Fisher
; TITLE OF INVENTION: "Plasmodium falciparum Ribonu-
; TITLE OF INVENTION: cleotide Reductase, DNA Sequences Therefor and Peptide Inhibi-
; TITLE OF INVENTION: Thereof"


```

; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The University of Pennsylvania
; STREET: Suite 330
; STREET: 3700 Market Street
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19104-3246
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 KB
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/136,743B
; FILING DATE: 10/14/93
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 3957-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5459063e
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 322 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-136-743B-2

Alignment Scores:
Pred. No.:          3,6e-108           Length:      322
Score:              974.50            Matches:     191
Percent Similarity: 75.62%             Conservative: 51
Best Local Similarity: 59.69%           Mismatches:   75
Query Match:        52.79%             Indels:       3
DB:                  1                 Gaps:         2

US-10-698-228-12 (1-1053) x US-08-136-743B-2 (1-322)

QY 100 CTCCTAAGAAAGATTCTCGCGGTTGTGCATCTTTCCAATCCAGTACCCTCATATTGG 159
Db 4 IleLeuAsnLysGluSerAspArgPheThrLeuTyProIleLeuTyProaspValPhe 23
QY 160 AAATGTATAACAGGCCACAGGCTTCCTCTCGGACAGCAGAAGAGTTGACTTATCAAAG 219
Db 24 ProPheTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluLeaAspTyrSerSer 43
QY 220 GATCTCCCTCACTGGAACAAGCTTAAGACAGATGAGAAGTAGTCTTCATCTCTCACATCTTA 279
Db 44 AspLeuLysAspPheGluLysLeuAsnGluAsnGluLysHisPheIleLysHisValLeu 63
QY 280 GCCTTTTTTGCCAGCCAGTGATGGAATGTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAG 339
Db 64 AlaPhePheAlaAlaSerAspDlyIleValLeuGluAsnLeuAlaValSerPheLeuArg 83
QY 340 GAGGTGCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTTCTCATCGAAGATGTT 399
Db 84 GluValGlnIleThrGluAlaLysLysPheTySerPheGlnIleAlaValGluAsnIle 103
QY 400 CACTTCAGAGATGTACAGTTTGTGTATGACACACTTACATCAGAGATCCCAGAAAAAGGGAA 459
Db 104 HisserGluThrTySerLeuLeuIleAspAsnTyrlleLysAspGluLysGluArgLeu 123
QY 460 TTTTATTATTAATGCAATTGAACCATGCCCTATGTTAAGAAAAAACGAGATTGGCGCTTGG 519
Db 124 AsnLeuPheHisAlaIleGluAsnIleProAlaValLysAsnLysAlaLeuTrpAlaAla 143
QY 520 CGATGGATACGATAGAAAAATCTACTTTTGGGGAAACAGTGTGGCGCTTGTGCTGCTGTA 579

```

```
QY 127 GTCACTTTCCAAATCCAGTACCTGATATTGTGAAAATGTATAAACAGGCAGGCTTCC 186
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
112 ValMetPheProIleArgTyrHisGluIleTyrAsnPhetYrIysIysAlaGluAlaSer 131
QY 187 TTCTGGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCAGTGG---AACAGCTT 243
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
132 PheTyrThrAlaGluIleAspLeuSerIysAspLeuAspTrpAsnAsnIysLeu 151
QY 244 AAAGCAGATGAGAAGTACTTCTCTCATCTTACGCTTTTTCAGCCAGTGTGGA 303
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
152 AsnGluAsnGluArgTyrPheIleSerArgValLeuAlaPheAlaSerAspGly 171
QY 304 ATTGTAAATGAAAATTTGGTGGAGCCTTTAGTCAGAGGTGAGTTCAGAGGCTCGC 363
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
172 IleValGlyLeuAsnLeuIleGluAsnPheSerThrGluValGluLeuProGluAlaIys 191
QY 364 TGTCTTATGCGTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGACAGTTTGCTG 423
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
192 SerPheTyrGlyPheGlnIleMetMetGluAsnIleHisSerGluThrTyrSerLeuLeu 211
QY 424 ATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATGAAACC 483
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
212 IleGluThrTyrIleIysAspProGlnGluAlaAspTyrLeuPheAsnAlaIleAsn 231
QY 484 ATGCCCTATGTTAAGAAAAAGCAGATTGGCGCTTCGGATGGATGAGATAGAAAATCT 543
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
232 IleProCysIleGlnIysAlaAspTrpAlaIleIysTrpIleGlnAspGluAla 251
QY 544 ACTTTTGGGGAAGAGTGGTGGCTTTGCTGTGTAGAAGGATTTTCTTCAGGATCT 603
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
252 LeuTyrGlyGluArgLeuValAlaPheAlaAlaValGluGlyIlePhePheSerGlySer 271
QY 604 TTGTCTGCTATATTCGGCTAAAGAGAGAGGTCTTATGCCAGGATCTCTTTTCCAAAT 663
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
272 PheAlaSerIlePheTyrLeuIysIysArgGlyLeuMetProGlyLeuThrPheSerAsn 291
QY 664 GAACCTCATCAGCAGAGATCAAGGATCTCACTGTGACTTTGCTGTGCTGATGTTCCATAC 723
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
292 GluLeuIleCysArgAspGluGlyLeuHisThrAspPheAlaCysLeuLeuPheAlaHis 311
QY 724 TTAGTAAATAAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTCGTCAAAAT 783
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
312 LeuGlnAsnArgProSerProGluIleValGluArgIleIleThrGluAlaValAspVal 331
QY 784 GAGCAGGAGTTTAAAGAGCCTTGGCAGTTGGCTCATTTGGAATGATTCATTG 843
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
332 GluLysGluTyrPheThrAspValLeuProValSerLeuLeuGlyLeuAsnCysAsnLeu 351
QY 844 ATGAACAGTACATTCAGTTGTA 867
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
352 MetCysProTyrIleGluPheLeu 359
```

RESULT 10

```
US-08-905-223-358
; Sequence 358, Application US/08905223
; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 358:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 147 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Brain
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: -101...-1
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 4.4
; OTHER INFORMATION: seq ISHILAFFAASDG/IV
; US-08-905-223-358
```

Alignment Scores:

Pred. No.:	1,75e-83	Length:	147
Score:	768.00	Matches:	147
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	41.60%	Indels:	0
DB:	3	Gaps:	0

US-10-698-228-12 (1-1053) x US-08-905-223-358 (1-147)

QY	1	ATGGGCGACCGGAAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA	60
Db	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer	20
QY	61	GACACCAACAAAGTGAATAAAGTCAATGAAGAGCCACTCCTAAGAAAGATTCTCGC	120
Db	21	AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg	40
QY	121	CGGTTTGTCACTTCCCAATCCAGTACCCTGATATTTGGAAAATGTATAAACAGGCACAG	180
Db	41	ArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysGlnAlaGln	60
QY	181	GCTTCTCTTCGACAGCAGAGAGGTTGACTTATCAAAAGGATCTCCCTCAGTGAACAG	240
Db	61	AlaSerPheTyrThrAlaGluValAspLeuSerLysAspLeuProHisTyrAsnLys	80
QY	241	CTTAAAGCAGATCAGAAAGTACTTCATCTCACAATCTAGCCTTTTTCGCGCCAGTGTAT	300
Db	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaSerAsp	100
QY	301	GGAATTCTAAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCCAGGCTCCAGGCT	360
Db	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla	120
QY	361	CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG	420
Db	121	ArgCysPheTyrGlyPheGlnIleLeuIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140
QY	421	CTCATGACACTTACATCAGA 441	
Db	141	LeuIleAspThrTyrIleArg 147	

RESULT 11

```
US-09-270-767-46236
; Sequence 46236, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46236
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46236

Alignment Scores:
Pred. No.: 3,98e-63 Length: 149
Score: 599.00 Matches: 114
Percent Similarity: 85.91% Conservative: 14
Best Local Similarity: 76.51% Mismatches: 21
Query Match: 32.45% Indels: 0
DB: 4 Gaps: 0

US-10-698-228-12 (1-1053) x US-09-270-767-46236 (1-149)
QY 586 GTTTCCTCTCAGGATCTTTCGCTGCTATATCTGCTAAAGAGAGAGGTCTTATGCCA 645
Db 1 IIPhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArgGlyLeuMetPro 20
QY 646 GGACTCACCTTTCCAAATCAATCATCAGCAGAGAGGAGGACTTCACTGTGACTTTGCT 705
Db 21 GlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAla 40
QY 706 TGCCTGATCTTCCAACTACTAGTAATAAGCCCTTCAGAGAAAGGCTCAGGAGATCAAT 765
Db 41 ValLeuMetPheGlnHisLeuValGlnArgProLysArgGluArgIleLeuGluIle 60
QY 766 GTTGATGCTGCAAAATGAGCAGGAGTCTTTTAAACAGAGCCCTGCGAGTGGCCCTCAT 825
Db 61 ArgAspAlaValAlaIleGluGlnPheLeuThrAspAlaLeuProValAsnLeuIle 80
QY 826 GCAATGAATGCAATTTGATGAACAGTACATGAGTTGTAGCTCAGATTAATCTTGTG 885
Db 81 GlyMetAsnCysAspLeuMetSerGlnTyrlleGluPheValAlaAspArgLeuLeuVal 100
QY 886 GAACCTGGATCTCAAGGTTTTCAGGCAGAGAAATCCCTTTTATGTTTATGAAACAT 945
Db 101 GluLeuGlyValGlyLysIleTyrlAsnThrLysAsnProPheAsnPheMetGluMetIle 120
QY 946 TCTTTAGAGGAAACAAATTTCTTTGAGAACCGATTTTCAGAGTATCAGCGCTTTTGCA 1005
Db 121 SerLeuAspGlyLysThrAsnPhePheGluLysLysValGlyGluTyrlGlnArgMetGly 140
QY 1006 GTTATGGCAGAACACAGATACGTC 1032
Db 141 ValValSerAsnProLeuAspAsnVal 149

RESULT 12
US-09-902-540-10006
; Sequence 10006, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
```

```
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10006
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10006

Alignment Scores:
Pred. No.: 2.01e-50 Length: 323
Score: 495.00 Matches: 113
Percent Similarity: 54.49% Conservative: 63
Best Local Similarity: 34.98% Mismatches: 125
Query Match: 26.81% Indels: 22
DB: 4 Gaps: 6

US-10-698-228-12 (1-1053) x US-09-902-540-10006 (1-323)
QY 145 TACCCTGATATTTGAAAAATGTATAAACAGCAGCAGCGCTTCTCTTGTGACAGCAGAGAG 204
Db 3 TyrProGlnPhePheGluMetTyArgAsnAlaIleLysAsnThrTrpThrValGluGlu 22
QY 205 GTTGACTATCAAGAGGATCTCCCTCACTGG---AACAGCTTAAAGCAGATGAGAGATAC 261
Db 23 ValAspPheSerThrAspLeuValAspLeuArgSerLysMetThrAspAlaGluArgHis 42
QY 262 TTCATCTCTCACATCTTAGCCTTTTTCGACCCAGTGTATGGAATTTGTAATGAAATTTG 321
Db 43 LeuIleHisArgLeuValAlaPhePheAlaThrGlyAspSerIleValGlyAsnAsnLeu 62
QY 322 GTCGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGCGCTCGCTGTTCTATGCTTCA 381
Db 63 ValLeuAsnLeuTyrlLysHisLeuAsnAlaProGluAlaArgMetTyrlLeuSerArgGln 82
QY 382 ATTCTCATCGAGAATGTTCATCTCAGAGATGTACAGTTTGTGTATAGACACTTACATCAGA 441
Db 83 LeuTyrlGluGluAlaLeuHisValGlnPheTyrlLeuThrLeuLeuAspThrTyrlValPro 102
QY 442 GATCCCAAGAAAAGGGAATTTTATTAATGCAATTTGAAACCATGCTTATGTTTAAAGAAA 501
Db 103 AspProAlaGluArgAlaMetAlaPheAlaValAspAsnIleProSerIleGlnArg 122
QY 502 AAAGCAGATTTGGCGCTTCCGATGATA-----GCAGATAGAAAATCTACT 546
Db 123 LysAlaGlnPheCysMetLysTrpMetAspSerIleGlnGlyLeuAspThrLeuGlnThr 142
QY 547 TTTGGGGAAGA-----GTGTGGCTTTTGCT---GCTGTAGAGGA 585
Db 143 LysAlaHisArgArgGlnPheLeuLeuAsnLeuIleCysPheAlaGlyCysIleGluGly 162
QY 586 GTTTCCTCTCAGAGATCTTTTGTGCTATATCTGGCTAAAGAGAGAGGTCTTATGCCA 645
Db 163 LeuPhePhePheAlaAlaPheAlaTyrlValTyrlPheLeuArgSerLysGlyLeuLeuAsn 182
QY 646 GGACTCACTTTTCCCAATGACATCAGCAGAGATGAAGACTTTCAGCTGAGCTTGTCT 705
Db 183 GlyLeuAlaAlaGlyThrAsnTrpValPheArgAspGluSerAlaHisMetAlaPheAla 202
QY 706 TGCCTGATGTTCCAATAC-----TTAGTAAATAAGCCTTCAGAA 744
Db 203 PheGluSerIleGlnValAlaArgLysGluGluProAspLeuPheAspAlaGlnMetGlu 222
QY 745 GAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTTGACGAGGAGTCTTAAACAGAA 804
Db 223 ArgAspValValAlaMetLeuArgGluAlaValGluCysGluThrGlnPheAlaGlnAsp 242
QY 805 GCCTTGCAGTTCGCTCATTTGAATGAATTTGATGAAACACAGTACATGAGTTT 864
Db 243 LeuLeuSerGlyGlyValMetGlyLeuSerValGlnAspMetArgGlyTyrlLeuGluTyrl 262
QY 865 GTAGCTGACAGATTAATCTTGTGGAACCTTGGATTTCTCAAGAGTCTTTCAGGCGAGAAATCCT 924
Db 263 ValAlaAspGlnArgLeuGlnMetLeuGlyMetThrProValPheGlyThrLysAsnPro 282
```



```
Db 99 GlySerValGluGluAlaAlaValSerAlaAspSerAspAlaValAlaAlaArgAla 118
QY 55 TCTTCAGACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCCTCTCAAGAAAGT 114
Db 119 LysLysAlaLeuAsnAspLeuAspIleGlnGluGlyLeuAspAspLeuGluGlySerAla 138
QY 115 TCTCGCGGTTTGTC-----ATC 132
Db 139 AlaArgValGlnValGlyAspLysGlnMetIleAsnAlaArgAlaAspLeuAsnGlnLeu 158
QY 133 TTTCCAATCCAGTACCTGATATTGGAAAAATGATAAAGCAGCAGGCTTCCTCTCGG 192
Db 159 ValProPheLysTyAspTrpAlaTrpGlnLysTyLeuAspGlyCysAlaAsnHisTrp 178
QY 193 ACAGCAGAGAGGTTGACTTATCAAGAGTCTCCCTCACTGGNACAAG-----CTT 243
Db 179 MetProGlnGluValAsnMetAsnAlaAspIleAlaLeuTrpLysSerLysAspGlyLeu 198
QY 244 AAAGCAGATGAAGTACTTCACTCTCACATCTTAGCCCTTTTTCAGCCAGTGATGGA 303
Db 199 SerGluHisGluArgIleValMetArgAsnLeuGlyPhePheSerThrAlaAspSer 218
QY 304 ATTGTAATGAATAATTGGTGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCTCGC 363
Db 219 LeuValAlaAsnAsnLeuValLeuAlaValTyArgLeuIleThrAsnProGluCysArg 238
QY 364 TGTTCCTATGGCTTCAAAATTCATCGAGAATGTTCACTCAGAGATGACAGTTTGTG 423
Db 239 GlnTyIleLeuArgGlnAlaPheGluGluAlaIleHisThrHisAlaTyArgLntyrCys 258
QY 424 ATAGACACTTACATCAGAGATCCCAAGAAAGGAATTTTATTAATGCAATTGAAACC 483
Db 259 IleGluSerLeuGlyMetAspGluGlyGlu-----IlePheAsnMetTyrHisGlu 275
QY 484 ATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTGGCATGATA----- 528
Db 276 IleProSerValAlaLysLysAlaSerTrpGlyLeuLysTyThrArgSerIleSerAsp 295
QY 529 -----GCAGATAGAAAATCTACTTTTGGGGAAGAGTGG 561
Db 296 ProMetPheGlnThrGlyThrProGluThrAspArgGln-----PheLeuArgAsnLeu 313
QY 562 GTGGCTTTGTGCTGTGA---CAAGAGTTCCTCTCAGGATCTTTTCTGCTATATTC 618
Db 314 IleAlaTyTrpCysValLeuGluGlyIlePhePheTyCysGlyPheThrGlnIleLeu 333
QY 619 TGGCTAAGAAGACAGAGGTTCTTATGCCAGACTCACTTTTCCAAATCACTCAGCAGA 678
Db 334 SerMetGlyArgArgAsnLysMetThrGlyThrAlaGluGlnPheGlnTyIleLeuArg 353
QY 679 GATGAAGGACTTCATGTGACTTTCCTGTC---CTGATGTTCCAACTACTAGTAATAAG 735
Db 354 AspGluSerMetHisLeuAsnMetGlyIleAspValIleAsnGlnIleLysIleGluAsn 373
QY 736 CCT-----TCAGAAGAAAGGTCAGGAGATCATTTGCTGATCTGTC 777
Db 374 ProHisLeuTrpAspAlaGlnMetLysAspGluAlaThrGlnMetIleLeuGlnGlyThr 393
QY 778 AAAATTGACAGGAGGTTTTTAACAGAGCTTGGCAGTTGGCTCATTTGGAATGAATTGC 837
Db 394 GlnLeuGluIleGluTyAlaArgAspThrMetProArgGlyValLeuGlyMetAsnAla 413
QY 838 ATTTTGATGAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGGAATCTGGATTC 897
Db 414 AlaMetGluAspTyLeuLysPheIleAlaAsnArgLeuThrGlnIleGlyLeu 433
QY 898 TCAGAAGTTTTTCAG---GCAGAAATCTTTTGAATTTATGGAATAACATT---TCCTTA 951
Db 434 LysGluGluTyProGlyThrThrAsnProPheProTrpMetSerGluIleMetAspLeu 453
QY 952 GAAGGAAAAACAAATTTCTTTTGAGAAACAGATTTTCAGAGTATCAG 996
```

```
Db 454 LysLysGluLysAsnPhePheGluThrArgValIleGluTyArgLntyrGln 468
RESULT 15
US-09-198-452A-1059
; Sequence 1059, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1059
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-1059

Alignment Scores:
Pred. No.: 3,98e-31 Length: 365
Score: 335.00 Matches: 84
Percent Similarity: 46.84% Conservative: 64
Best Local Similarity: 26.58% Mismatches: 134
Query Match: 18.15% Indels: 34
DB: 4 Gaps: 6

US-10-698-228-12 (1-1053) x US-09-198-452A-1059 (1-365)
QY 130 ATCTTTCCAAATCCAGTACCTGATATTTGGAAAATGTATAACAGGCACAGGCTTCCTTC 189
Db 50 LeuValProIleLysTyArgTrpAlaTrpGluHisTyLeuAsnGlyCysAlaAsnAsn 69
QY 190 TGCAGCAGCAGAGAGGTTGACTTATCAAGAGTCTCCCTCACTGG-----AACAAAGCTT 243
Db 70 TrpLeuProThrGluValProMetAlaArgAspIleGluLeuTrpLysSerAspGluLeu 89
QY 244 AAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCTTTTTCAGCCAGCAGGATGAGA 303
Db 90 SerGluAspGluArgValIleLeuLeuAsnLeuGlyPhePheSerThrAlaGluSer 109
QY 304 ATTGTAATGAATAATTGGTGAGCGCTTTAGTCAGGAGGTCAGGTTCCAGAGGCTCGC 363
Db 110 LeuValGlyAsnAsnIleValLeuAlaIlePheLysHisIleThrAsnProGluAlaArg 129
QY 364 TGTTCCTATGGCTTCAAAATTTCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTGTG 423
Db 130 GlnTyIleLeuArgGlnAlaPheGluGluAlaValHisThrHisThrPheLeuTyIle 149
QY 424 ATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAAACC 483
Db 150 CysGluSerLeuGlyLeuAspGluGlyGlu-----ValPheAsnAlaTyArgAsnGlu 166
QY 484 ATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTGGCATGGAATGATGATGATGATGATGAT 534
Db 167 ArgAlaSerIleArgAlaLysAspPheGlnMetThrLeuThrValAspValLeuAsp 186
QY 535 -----AGAAAATCTACTTTTGGG 552
Db 187 ProAsnPheSerValGlnSerSerGluGlyLeuGlyGlnPheIleLysAsnLeuValGly 206
QY 553 GAAAGAGTGGTGGCCCTTGTGCTGTAGAGAGGAGTTTCTTCTCAGGATCTTTTGTGCTGT 612
Db 207 TyrTyIleIle-----MetGluGlyIlePhePheTySerGlyPheValMet 222
QY 613 ATATTCGTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTCCAAATGAACTCATC 672
Db 223 IleLeuSerPheHisArgGlnAsnLysMetThrGlyIleGlyGluGlnTyArgLntyrIle 242
QY 673 AGCAGAGATGAAGGACTTCACTGTGACTTTGTGCTGTGATGTTCCCAATACTTAGTAAT 732
```

```
Db      243 LeuArgAspGluThrIleHisLeuAsnPheGlyIleAspLeuIleAsnGlyIleLysGlu 262
QY      733 AAGCCTTCAGAA-----GAAAGGTCAGGAGATCATTGTTGAT 771
Db      263 GluAsnProGluValTrpThrGluLeuGlnGluIleValAlaLeuIleGluLys 282
QY      772 GCTGTCAAAATTGAGCAGGAGTTTTAAACAGAAGCCTGCCAGTTGGCCCTCATTGGAATG 831
Db      283 AlaValGluLeuGluIleGluTyralaLysAspCysLeuProArgGlyIleLeuGlyLeu 302
QY      832 AATTGCATTTTGATGAACACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGGAACTT 891
Db      303 ArgSerSerMetPheIleAspTyrValArgHisIleAlaAspArgLeuGluArgIle 322
QY      892 GGATTCTCAAAGGTTTTCAGGCAGAAATCCTTTTGATTATG--GAAAACATTCT 948
Db      323 GlyLeuLysProIleTyrHisSerArgAsnProPheProTrpMetSerGluThrMetAsp 342
QY      949 TTAGAAGGAAACAAATTTCTTGAGAAACGAGTTTCAGAGTATCAG 996
Db      343 LeuAsnLysGluLysAsnPheGluThrArgValThrGluTyrGln 358
```

Search completed: October 30, 2005, 06:36:21
Job time : 24.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 30, 2005, 06:34:06 ; Search time 65.75 Seconds
(without alignments)
13390.093 Million cell updates/sec

Title: US-10-698-228-12

Perfect score: 1846

Sequence: 1 atggcgaccgcaaggcc.....tcaccttgatgcagattttt 1053

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1865214 seqs, 418043040 residues

Total number of hits satisfying chosen parameters: 3730428

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10698228/runat_26102005_100610_4101/app_query.fasta_1.2446
-DB=Published Applications AA -QWTF=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEARSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USBR=US10698228 @cgn 1 113 @runat_26102005_100610_4101
-NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FPGAPOP=6 -FPGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10F_PUBCOMB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1821	98.6	351	15	US-10-698-228-1	Sequence 1, Appli
2	1454	78.8	389	15	US-10-698-228-5	Sequence 5, Appli
3	1454	78.8	389	16	US-10-408-765A-559	Sequence 559, App
4	1454	78.8	389	16	US-10-733-878-457	Sequence 457, App
5	1454	78.8	389	18	US-10-756-149-4899	Sequence 4899, Ap
6	1454	78.8	413	9	US-09-925-301-1347	Sequence 1347, Ap
7	1454	78.8	453	18	US-10-220-335-686	Sequence 686, App
8	1454	78.8	453	18	US-10-450-763-45646	Sequence 45646, A
9	1425.5	77.2	386	15	US-10-403-571-76	Sequence 76, Appl
10	1348	73.0	393	20	US-11-097-143-19182	Sequence 19182, A
11	1242	67.3	430	15	US-10-369-493-4122	Sequence 4122, Ap
12	1230	66.6	405	14	US-10-128-714-8245	Sequence 8245, Ap
13	1214	65.8	381	15	US-10-369-493-5686	Sequence 5686, Ap
14	1192.5	64.6	381	14	US-10-128-714-3245	Sequence 3245, Ap
15	1192	64.6	413	14	US-10-032-585-7537	Sequence 7537, Ap
16	1176.5	63.7	321	15	US-10-369-493-10365	Sequence 10365, A
17	1169.5	63.4	399	15	US-10-369-493-22170	Sequence 22170, A
18	1161.5	62.9	391	15	US-10-369-493-22565	Sequence 22565, A
19	1161.5	62.9	401	15	US-10-320-797-3145	Sequence 3145, Ap
20	1147.5	62.2	329	15	US-10-424-599-215866	Sequence 215866,
21	1133.5	61.4	717	16	US-10-425-115-215845	Sequence 215845,
22	1132.5	61.3	345	16	US-10-437-963-113592	Sequence 113592,
23	1128.5	61.1	339	16	US-10-767-701-45287	Sequence 45287, A
24	1125.5	61.0	339	16	US-10-425-115-247984	Sequence 247984,
25	1114.5	60.4	329	15	US-10-424-599-269685	Sequence 269685,
26	1097.5	59.5	339	15	US-10-424-599-183313	Sequence 183313,
27	1092.5	59.2	400	16	US-10-425-115-200125	Sequence 200125,
28	1061	57.5	347	16	US-10-739-930-6922	Sequence 6922, Ap
29	916	49.6	270	18	US-10-220-335-342	Sequence 342, App
30	871.5	47.2	344	15	US-10-369-493-13401	Sequence 13401, A
31	840	45.5	254	16	US-10-425-115-215843	Sequence 215843,
32	805	43.6	345	15	US-10-369-493-21987	Sequence 21987, A
33	784	42.5	225	16	US-10-425-115-215840	Sequence 215840,
34	701	38.0	138	15	US-10-264-237-2195	Sequence 2195, Ap
35	680	36.8	335	18	US-10-450-763-45645	Sequence 45645, A
36	662	35.9	181	18	US-10-450-763-37219	Sequence 37219, A
37	564.5	30.6	338	14	US-10-156-761-10564	Sequence 10564, A
38	544	29.5	156	15	US-10-424-599-215870	Sequence 215870,
39	491	26.6	307	15	US-10-369-493-19363	Sequence 19363, A
40	456.5	24.7	145	16	US-10-767-701-45286	Sequence 45286, A
41	397	21.5	398	15	US-10-282-122A-51357	Sequence 51357, A
42	384.5	20.8	415	15	US-10-282-122A-66252	Sequence 66252, A
43	384	20.8	403	15	US-10-282-122A-50118	Sequence 50118, A
44	383.5	20.8	415	15	US-10-282-122A-69814	Sequence 69814, A
45	380.5	20.6	416	15	US-10-282-122A-68043	Sequence 68043, A

ALIGNMENTS

RESULT 1

US-10-698-228-1
; Sequence 1, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/698, 228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 1

```
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-1
Alignment Scores:
Pred. No.: 4.54e-187 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 15 Gaps: 0
US-10-698-228-12 (1-1053) x US-10-698-228-1 (1-351)
QY 1 ATGGGCGACCCGGAAGGCGGAGCGCGGCTGGATCAGATGAGAGATCAATCTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAAGCAAGTAAAGTCAATGAGAGCCACTCTTAAAGAGTCTCCTC 120
Db 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
QY 121 CGGTTTGTCTTCCCAATCCAGTACCCTGATATTGGAATGTAATAACAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTCTGACAGCAGCAAGAGGTGACTTATCAAGGATCTCCTCCTCACTGGAACAAG 240
Db 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuPheHisTrpAsnLys 80
QY 241 CTTAAGCAGATGAGAGTACTCTCTCATCTCTACATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GCAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTCCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTATCGAGAATGTTCTCACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTATCATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTGGCCCTTCGATGGATAGCAGATAGAAA 540
Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTAGAGAGGAGTCTTCTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTGCTATATCTGGCTTAAAGAGAGAGGTCTTATGCGAGACTCACTTTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 ATGAACCTCATCAGCAGATGAGAGCTTCACTGTGACTTTGCTTGGCTGCTGATGTTCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAGCTTTCAGAGAAAGGTCAGGAGATCATCTGTGATGCTGTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGAGCAGAGTTTTTAACAGAGCGCTTGCAGAGTGGCTCATTTGGAATGTAATTCATC 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAAACAGTACATTTGTTAGTCTGACAGATTTACTTGTGGAACCTTGTGATTTCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGTTTTTTCAGGCAGAAAATCCTTTTGTGATTTTATGAAAACATTTCTTTAGAGAGAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTTCAGAAACAGTTTTCAGAGTATTCAGCGTTTTTGCAGTTATTCGCAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTTTCACCTTTGGATGCAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
RESULT 2
US-10-698-228-5
; Sequence 5, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 5
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-5
Alignment Scores:
Pred. No.: 1.96e-147 Length: 389
Score: 1454.00 Matches: 389
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 15 Gaps: 0
US-10-698-228-12 (1-1053) x US-10-698-228-5 (1-389)
QY 91 GAAGAGCCACTCTCTAAGAAAGAGTCTTCGCGGTTTGTTCATCTTTCCAATCCAGTACCT 150
Db 69 AspGluProLeuLeuArgGluAsnProArgArgPheValIlePheProIleGluTyrHis 88
QY 151 GATATTGGAAATGATATAAACAGGCAGCTTCTTCTGGAACAGAGAGAGGTTCAC 210
Db 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
QY 211 TTATCAAGAGATCTCCTCACTGGAACAAGCTTAAAGCAGATGAGAGTACTTCTCTCT 270
Db 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY 271 CATCTTACCTCTTTTTCAGCCAGGTGATGGAAATGTAATGAAATTTTGGTGGAGCGC 330
Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY 331 TTTAGTCAGGAGGTGCGAGTCCAGAGGCTCGCTGTTTCTATGCTTTTCAATTTCTCATC 390
Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
```


Db 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
QY 211 TTATCAAGAGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATCAGAAAGTACTTTCATCTCT 270
Db 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTTAGCCCTTTTTCAGCCAGTGTGGAATTTGTAATGAAATTTGGTGGAGCGC 330
Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAlaValGluArg 148
QY 331 TTATAGTCAGAGGTGTCAGGTTCAGAGGCTCGCTGTTCTATGGCTTCAAAATTCATC 390
Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGAAATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCAAG 450
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAAGGGAATTTTATTTAATGCAATTGAACCATGCCCTATGTTAAGAAAAAGCAGAT 510
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
QY 511 TGGCCCTTCGGATGATAGATGATGAAATCTACTTTTGGGAAAGAGTGTGGCTTTT 570
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 228
QY 571 GCTGCTGTAGAGGAGTTTCTTCTCAGAGTCTTGTGCTGCTATATCTGGCTAAAGAG 630
Db 229 AlaAlaValGluGlyIlePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGGTCTTATGTCAGGACTCTACTTTTCCAAATGAACTCATCAGCAGAGATGAAGACTT 690
Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY 691 CACTGTGACTTGTGCTGCTGATGTTTCCAAATCTTGTAAATTAAGCCTTCAGAGAAAGG 750
Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGAGATCATTTGTCATGCTGTCAAAATGACAGAGTCTTAAACAGAGCCTTG 810
Db 289 ValArgGluIleIleLeuAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCCTCATTTGGAATGAATTCATTTTGAATGAAACAGTACATTCAGTTTGTAGCT 870
Db 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
QY 871 GACAGATTACTTGTGGAATTCGATTTCTCAAAAGGTTTTTTCAGGCAAAAATCCTTTGAT 930
Db 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAlaSerProPheAsp 348
QY 931 TTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAGAG 990
Db 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY 991 TATCAGCGTTTTCAGTTATGTCAGAGAACCAACAGATAACGTTCTACCTTGGATGAGAT 1050
Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY 1051 TTT 1053
Db 389 Phe 389

RESULT 6

US-09-925-301-1347
; Sequence 1347, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882

; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1347
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1347

Alignment Scores:
Pred. No.: 2,01e-147 Length: 413
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: Gaps: 0

US-10-698-228-12 (1-1053) x US-09-925-301-1347 (1-413)

QY 91 GAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTTGTCACTTTCCTCAATCAGTACCCT 150
Db 93 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 112
QY 151 GATATTTGGAAATGTATAACAGGCACAGGCTTCTCTGACAGCAGAGAGGCTTGCAC 210
Db 113 AsplleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 132
QY 211 TTATCAAGAGTCTCCCTCACTCGAACAAGCTTAAAGCAGATGAGAACTTCTCATCTCT 270
Db 133 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 152
QY 271 CACATCTTAGCCCTTTTTCAGCCAGTGTGGAATTTGTAATGAAATTTGGTGGAGCGC 330
Db 153 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAlaValGluArg 172
QY 331 TTATAGTCAGAGGTGTCAGGTTCAGAGGCTCGCTTCTATGGCTTCAAAATTCATC 390
Db 173 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 192
QY 391 GAGAAATGTTTATTTAATGCAAAATTCATGCTGCTGCTATATCTGCTGCTGCTGCTGCT 450
Db 193 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 212
QY 451 AAAAGGGAATTTTATTTAATGCAAAATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 510
Db 213 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 232
QY 511 TGGCCCTTCGGATGATGATGAAATTCATCTTTTGGGAAAGAGTGTGGCTTTT 570
Db 233 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 252
QY 571 GCTGCTGTAGAGGAGTTTCTTCTCAGAGTCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 630
Db 253 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 272
QY 631 AGAGGTCTTATGTCAGGACTCAGCTTTTCCAAATGAACTCATCAGCAGAGATGAAGACTT 690
Db 273 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 292
QY 691 CACTGTGACTTGTGCTGCTGATGTTCCAAATGAACTTCAATGAAATGAACTTCAAGAAAGG 750
Db 293 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 312
QY 751 GTCAGGAGATCATTTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 810
Db 313 ValArgGluIleIleLeuAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 332
QY 811 CCAGTTGGCCTCATTTGGAATGAATTCATTTTGAATGAAACAGTACATTCAGTTTGTAGCT 870
Db 333 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 352

LOCATION: (148)...(203)
OTHER INFORMATION: Ribonucleotide reductase small subunit proteins domain
OTHER INFORMATION: identified by eMATRIX, accession number BL00368A, p-value=1.000e-
OTHER INFORMATION: 40, raw score of 36.98
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (167)...(440)
OTHER INFORMATION: Ribonucleotide reductases domain identified by PFAM,
OTHER INFORMATION: accession name ribonuc_red, E-value=9e-174, PFAM score of 507.8
US-10-450-763-45646

Alignment Scores:
Pred. No.: 2,08e-147 Length: 453
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 18 Gaps: 0

US-10-698-228-12 (1-1053) x US-10-450-763-45646 (1-453)

QY	91	GAAGACCACTCTAAGAAGAGTTCTCGCGGTTTGTCTATCTTCCATCCAGTACCCT	150
DB	133	AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis	152
QY	151	GATATTGGAAATGTATAACAGGCACAGGCTTCTCTGGACAGCAAGAGTTGAC	210
DB	153	AspIleTrpGlnMetTyrLysAlaGluAsePheTrpThrAlaGluGluValAla	172
QY	211	TTATCAAGGATCTCCCTCACTGGAAACAAGCTTAAACAGACAGTACGAGTACTCTCT	270
DB	173	LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer	192
QY	271	CACATCTAGCCTTTTTCAGCCAGTGTAGTAATGTAATGTAATGTAATGTAATGTA	330
DB	193	HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg	212
QY	331	TTTAGTCAGAGGTGAGGTTCAGAGGCTCGCTGTTCTATGCGCTTCAATCTCATC	390
DB	213	PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet	232
QY	391	GAGATGTTCCTCAGAGATGTACAGTTCGTGATACACACTTACATCAGAGATCCCAAG	450
DB	233	GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys	252
QY	451	AAAAGGAATTTTATTAATGCAATTGAAACCATGCCCTATGTTAAGAAAAAAGCAGAT	510
DB	253	GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp	272
QY	511	TGGGCTTCGGATGATAGACAGTAATAATCTACTTTTGGGAAAGAGTGGGCTTT	570
DB	273	TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaAlaPhe	292
QY	571	GCTGCTGTAGAAGAGTTCCTCTCAGGATCTTTTGTCTATATCTTGGCTTAAAGAG	630
DB	293	AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys	312
QY	631	AGAGTCTTATCCAGGATCTACTTTTCCAATGACTCATCAGCAGAGATGAAGACTT	690
DB	313	ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu	332
QY	691	CACCTGTGACTTTGCTGCTGATGTTCATACTTAGTAATAAGCTTCAGAAGAAAGG	750
DB	333	HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLeuValHisProSerGluGluArg	352
QY	751	GTACGGAGATCATTTGTTGATCTCTCAAAATTGACGAGGATTTTAAACAGAGCTTG	810
DB	353	ValArgGluIleIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu	372
QY	811	CCAGTTGGCCTCATTTGGAATGAAATTTTGATGAAACAGTACATTTGAGTTTGTAGCT	870
DB	373	ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla	392

QY	871	GACAGATTACTTGTGAACCTTGGATTCTCAAGAGTTTTCAGGCAGAAAAATCTTTGAT	930
DB	393	AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp	412
QY	931	TTATGGAAAAACATTTCTTTAGAGGAAAAACAATTTCTTTTGAGAAACGAGTTTCAGAG	990
DB	413	PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu	432
QY	991	TATCAGCGTTTTCAGATTATGGCAGAACACACAGATTAAGCTTTCACCTTGGATCGAGT	1050
DB	433	TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp	452
QY	1051	TTT 1053	
DB	453	Phe 453	

RESULT 9
US-10-403-571-76
Sequence 76, Application US/10403571
Publication No. US20040068763A1
GENERAL INFORMATION:
APPLICANT: Hopkins, Nancy
APPLICANT: Golling, Gregory
APPLICANT: Amsterdam, Adam
APPLICANT: Sun, Zhaoxia
TITLE OF INVENTION: Developmental Mutations in Zebrafish
FILE REFERENCE: 01997/539002
CURRENT APPLICATION NUMBER: US/10/403,571
CURRENT FILING DATE: 2003-03-25
PRIOR APPLICATION NUMBER: US 60/368,760
NUMBER OF SEQ ID NOS: 159
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 76
LENGTH: 386
TYPE: PRT
ORGANISM: Danio rerio
US-10-403-571-76

Alignment Scores:
Pred. No.: 2,34e-144 Length: 386
Score: 1425.50 Matches: 268
Percent Similarity: 89.41% Conservative: 36
Best Local Similarity: 78.82% Mismatches: 31
Query Match: 77.22% Indels: 5
DB: 15 Gaps: 1

US-10-698-228-12 (1-1053) x US-10-403-571-76 (1-386)

QY	49	AGATCATCTTCAGACACCAAGTGAATAAAGTCA-----AATGAA	93
DB	47	ArgLysIlePheAspGluSerGluGlyGlnSerLysAlaLysLysGlyAlaValGluGlu	66
QY	94	GAGCCACTCTAAGAAGAGTCTCGCGGTTTGTCTATCTTCCAAATCCAGTACCCTGAT	153
DB	67	GluProLeuLeuLysGluAsnProHisArgPheValIlePheProIleGlnTyrHisAsp	86
QY	154	ATTGTGAAATGTATAACAGGCACAGCTTCTCTGGACAGCAGAGAGGTTGACTTA	213
DB	87	IleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAspLeu	106
QY	214	TCAAAGATCTCCCTCAGTGGAAACAAGCTTAAAGCAGATGAGAGTACTTCTCATC	273
DB	107	SerLysAspLeuGlnHisTrpAspSerLeuLysAspGluGluArgTyrPheIleSerHis	126
QY	274	ATCTTAGCCTTTTTCAGCCAGTGAATTCGTAATGAATGAATTTGTCGAGCGCTTT	333
DB	127	ValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArgPhe	146
QY	334	AGTCAGAGGTGAGGTTCCAGAGGCTCGCTGTTTCTATGCTTTTCAAAATCTCATCGAG	393
DB	147	ThrGlnGluValGlnValThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMetGlu	166

```
QY 394 AATGTTCACTAGAGATGTACAGTTTGTGCTAGATAGACCTTACATCAGAGATCCCAAGAAA 453
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
167 AsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspSerLysGlu 186
QY 454 AGGGAATTTTATTAATCAATTGAACATGCGCTATGTTAAGAAAAAAGCAGATTGG 513
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
187 ArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAspTrp 206
QY 514 GCCTTTCGATGATGACAGATCAAAATCTACTTTTGGGGAAGAGTGGTGCCTTTGCT 573
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
207 AlaLeuAsnTrpIleGlyAspLysAsnAlaArgTyrGlyGluArgValValAlaPheAla 226
QY 574 GCTGTAGAAGAGTCTTCTTCTCAGAGATCTTTGCTGCTATATCTCGGCTAAGAGAGA 633
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
227 AlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArg 246
QY 634 GGTCTTATGCCAGGACTCACTTTTCCAAATGAACATCAGCAGAGATGAAGGACTTCAC 693
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
247 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHis 266
QY 694 TGTGATTTGCTGCTGATGTTCCAACTACTTAGTAATGAAGCTTCAGAGAAAGGCTC 753
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
267 CysAspPheAlaCysLeuMetPheLysHisLeuIleAsnLysProSerGluGluThrVal 286
QY 754 AGGAGATCATTTGTTGATGCTGCTCAAAATGAGCAGAGATTTTAAACAGAGCTTGCCA 813
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
287 LysLysIleIleMetAsnAlaValGluIleGluGluPheLeuThrAspAlaLeuPro 306
QY 814 GTTGCGCTCATTTGAATGAATTCATTTTGATGAAACAGTACATTCAGTTTGTAGCTGAC 873
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
307 ValLysLeuIleGlyMetAsnCysAspLeuMetLysGlnTyrIleGluPheValAlaAsp 326
QY 874 AGATTACTTGTGGAACTTGGATCTCAAGGTTTTCAGCGCAAAATCCTTTTGTATTT 933
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
327 ArgLeuLeuLeuGluLeuGlyPheAspLysValTyrArgValGluAsnProPheAspPhe 346
QY 934 ATGGAAACATTTCTTTAGAAGAAAAACAAATTTCTTTGAGAAACGAGTTTCAGAGTAT 993
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
347 MetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGluTyr 366
QY 994 CAGCGTTTTCAGTTATGCGCAAAACACAGATAACGCTTCACTTCCCTTGGATGCGATTT 1053
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
367 GlnArgMetGlyValMetSerGlyThrThrAspAsnThrPheThrLeuAspAlaAspPhe 386

RESULT 10
US-11-097-143-19182
; Sequence 19182, Application US/11097143
; Publication No. US2005020858A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
```

```
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19182
; LENGTH: 393
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-19182

Alignment Scores:
Pred. No.: 5,47e-136 Length: 393
Score: 1348.00 Matches: 252
Percent Similarity: 85.92% Conservative: 41
Best Local Similarity: 73.90% Mismatches: 48
Query Match: 73.02% Indels: 0
DB: 20 Gaps: 0

US-10-698-228-12 (1-1053) x US-11-097-143-19182 (1-393)
QY 31 GGGCTGGATCAGGATGAGAGATCATCTTCAGACACCAAGAAAGTGAATAAAGTCAAT 90
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
53 GlyIleGlyLysSerAlaAsnSerLeuMetGluLysSerValThrProPheAspProSer 72
QY 91 GAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTTCCAAATCCAGTACCCT 150
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
73 LeuGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGlnTyrHis 92
QY 151 GATATTGGAAAATGTATAACAGGCACAGCTTCTTCGACAGCAGAGAGAGGTTGAC 210
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
93 AspIleTrpGlnMetTyrLysAlaGluAlaSerPheTrpThrValGluGluValAsp 112
QY 211 TTATCAAAAGATCTCCCTCTCAGCAACAGCTTAAAGCAGATGAGAGTACTTCATCTCT 270
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
113 LeuSerLysAspLeuThrAspTrpHisArgLeuLysAspAspGluArgHisPheIleSer 132
QY 271 CACATCTTAGCCTTTTTCAGCCAGTGTGAATGTAATGAAATTTGGTGGAGCGC 330
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
133 HisValLeuAlaPhePheAlaIleSerAspGlyIleValAsnGluAsnLeuValGluArg 152
QY 331 TTTAGTCAGAGGTGCGAGTTCCAGAGGCTGCTGTTCTATGGCTTTCAATTCATC 390
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
153 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 172
QY 391 GAGAATGTTCACTCAGAGATGTACAGTTGCTGTAGTACACTTACATCAGAGATCCCAAG 450
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
173 GluAsnValHisSerGluMetTyrSerValLeuIleAspThrTyrIleArgAspProHis 192
QY 451 AAAAGGGAATTTTATTTAATGCAATGAAACCAATGCCCTATGTTAAGAAAAAAGCAGAT 510
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
193 GlnArgGluTyrLeuPheAsnAlaIleGluThrMetProAlaValLysArgLysAlaAsp 212
QY 511 TGGGCTTGGATGGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGGTGCCTTT 570
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
213 TrpAlaLeuSerTrpIleSerSerLysSerAlaAsnPheGlyGluArgIleIleAlaPhe 232
QY 571 GCTGCTGTAGAAGAGTTTCTCTCAGGATCTTTTGTGCTGTATATTCTGGCTAAAGAG 630
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
233 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 252
QY 631 AGAGGTCTTATGCCAGGACTCAGCTTTTCCAAATGAACTCATCAGCAGAGATGAAGACTT 690
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
253 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 272
QY 691 CACTGTGACTTTGCTTGCCTGATGTTCCAAATCATCTAGTAATAGACCTTCAGAAAGAG 750
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
273 HisCysAspPheAlaValLeuMetPheGlnHisLeuValGlnArgProLysArgGluArg 292
QY 751 GTCAGGAGATCATTTGTTGATGCTGCTCAAAATTCAGCAGAGGAGTTTAAACAGACCTTG 810
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
293 IleIleGluIleIleArgAspAlaValAlaIleGluGlnGluPheLeuThrAspAlaLeu 312
QY 811 CCAGTTGCGCTCATTTGGAATGAATTCATTTTGATGAAACAGTACATTCAGTTGTAGCT 870
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
313 ProValAsnLeuIleGlyMetAsnCysAspLeuMetSerGlnTyrIleGluPheValAla 332
```



```

; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8245
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-8245

Alignment Scores:
Pred. No.: 3,03e-123 Length: 405
Score: 1230.00 Matches: 234
Percent Similarity: 81.95% Conservative: 52
Best Local Similarity: 67.05% Mismatches: 49
Query Match: 66.63% Indels: 14
DB: 14 Gaps: 3

US-10-698-228-12 (1-1053) x US-10-128-714-8245 (1-405)

QY 46 GAGAGATCATCTTCAGACACCAAGAGTGAATAAGTCAATCAAGAGCCACTCTTA 105
DB 58 GluA1SerLysValAlaGlnSerIleLysGluLeuGluAlaAsn--GluProLeuLeu 76
QY 106 AGAAGAGTTCTCGCGGTTTCTCATCTTCCAAATCCAGTACCTCCATATTTGGAATG 165
DB 77 GlnGluAsnProHisArgPheValLeuPheProIleLysTyrHisGluIleTrpGlnMet 96
QY 166 TATAAACAGGACAGGCTTCCTCTCGGACAGAGAGGTTGACTTATCAAGAGATCTC 225
DB 97 TyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluIleAspLeuSerLysAspLeu 116
QY 226 CTTCACTGG---AACAAAGTTAAAGCAGATGAGAAGTACTTCTCATCTCTAGGCC 282
DB 117 HisAspTrpAsnAsnArgLeuAsnAspAspGluArgTyrPheIleSerHisValLeuAla 136
QY 283 TTTTTCAGCCAGTATGGAATTGTAATGAAAATTTGGTCGAGCGCTTTAGTCAGGAG 342
DB 137 PhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuLeuGluArgPheSerAsnGlu 156
QY 343 GTGAGGTTCCAGAGCGCTGCTTTCTATGCTTCAAAATCTCATCGAGATGTTTAC 402
DB 157 ValGlnValProGluAlaArgCysPheTyrGlyPheGlnIleMetIleGluAsnIleHis 176
QY 403 TCAGAGATGTACAGTTTGTGTATAGACATTCATCAGAGATCCCAAGAAAAGGGAATTT 462
DB 177 SerGluThrTyrSerLeuLeuIleAspThrTyrIleLysGluProLysGlnArgThrTyr 196
QY 463 TTATTATGCAATTTGAACCAATGCCCTATGTTAAGAAAAGCAGATTTGGCGCTTGCA 522
DB 197 LeuPheAspAlaIleAspThrValProCysIleArgLysAlaGlnTrpAlaMetArg 216
QY 523 TGGATAGCAGATAGAAAATCTACTTTTGGGAAAGAGTGTGGCTTCTGCTGTAGAA 582
DB 217 TrpIleSerAspLysGluSerThrPheAlaGlnArgLeuValAlaPheAlaAlaValGlu 236
QY 583 GGAGTTTCTCTCAGGATCTTTTGTCTATATCTTGCTAAAGAGAGAGTCTTATG 642
DB 237 GlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArgGlyLeuMet 256
QY 643 CCAGGACTCATTCTTTTCCAAATGAACTCATCAGCAGAGATGAAGGACTTCACCTGTGCTTT 702
DB 257 ProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisThrAspPhe 726

```

```

QY 703 GCTTGCTGATGTTTCCAATACTTAGTAATAAGCCTTCAGAAAGAAAGGTCAGGGAGATC 762
DB 277 AlaCysLeuLeuPheSerHisLeuAsnHisArgProAspLysLysValIleGluAspIle 296
QY 763 ATTGTTGATGCTGTCAAAATTTAGCAGGAGTATTTTAACAGAACCTTGGCCGCTC 822
DB 297 IleValGluAlaValAlaIleGluGlnGluPheLeuThrAspAlaLeuProValAlaLeu 316
QY 823 ATTGAATGAATGTCATTTTGATGAAACAGTACATTCAGTTTGTAGCTGCAGAGATTACTT 882
DB 317 LeuGlyMetAsnSerLysLeuMetCysGlnTyrIleGluPheValAlaAspArgLeuLeu 336
QY 883 GTGCACTTGGATTTCTCAAGGTTTTCAGGCGAATAATCCTTTTGATTTTATGCAAAAC 942
DB 337 ValAlaLeuGlyAsnLysTyrPheAsnSerProAsnProPheAspPheMetGluSer 356
QY 943 ATTCTTTAGAGAGAAAACAAATTTCTTTGAGAAAACGAGTTTCAGAGTATCAGCGTTT 1002
DB 357 IleSerLeuAlaGlyLysThrAsnPheGluLysArgValGlyAspTyrGlnLysAla 376
QY 1003 GCAGTTTATGCGCAGAAACC-----ACAGAT 1026
DB 377 GlyValMetAlaSerThrLysLysAspAlaSerGlnAspThrAlaLysAspAlaAsnAsp 396
QY 1027 AACGCTTTCACCTTGGATGCGAGATTTT 1053
DB 397 GlyGlyLeuCysPheAspGluAspPhe 405

RESULT 13
US-10-369-493-5686
; Sequence 5686, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5686
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5686

Alignment Scores:
Pred. No.: 1,58e-121 Length: 381
Score: 1214.00 Matches: 231
Percent Similarity: 81.10% Conservative: 48
Best Local Similarity: 67.15% Mismatches: 61
Query Match: 65.76% Indels: 4
DB: 15 Gaps: 2

US-10-698-228-12 (1-1053) x US-10-369-493-5686 (1-381)

QY 34 CTGGATCAGATCAGATCATCTTCAGAC-----ACCAACGAAGTGAATAAG--- 84
DB 38 ValAspGlnThrLysAlaAlaSerAlaGluGluThrAsnAsnGluSerGluValAsnGlu 57
QY 85 ---TCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTCCCATC 141
DB 58 LeuAspAlaAspGluProMetLeuGlnAspLeuAspAsnArgPheValIlePheProLeu 77
QY 142 CAGTACCTCATATTTGGAAAATGTATAAACAGCAGAGGCTTCTCTCTGACAGCAGAA 201
DB 78 LysHisHisAspIleTrpAsnPheTyrLysLysAlaValAlaSerPheTrpThrValGlu 97

```

```
QY 202 GAGTTGACTTATCAAAAGGATCTCCCTCTCTGGAACAAGCTTAAAGCAGATGAGAAATGAC 261
Db 98 GluValAspLeuGlyLysAspMetAsnAspTrpGluLysMetAsnGlyAspGluGlnTyr 117
QY 262 TTCATCTCTCATCTTAGCTTTTTCAGCCAGTGATGGAATTTGTAATGAATAATTTG 321
Db 118 PheIleSerArgIleuAlaPhePheAlaAAserAspGlyIleValAsnGluAsnLeu 137
QY 322 GTGGAGCGCTTAGTCAGGAGTGCGAGGTTCCAGAGGCTCGCTGTTCTCTATGCTTTCAA 381
Db 138 CysGluArgPheSerAsnGluValGlnValSerGluAlaArgPhePheTyrGlyPheGln 157
QY 382 ATTCTCATCGAGAATGTTTCACATCAGAGATGTACAGTTGCTGATGACACTTACATCAGA 441
Db 158 IleAlaIleGluAsnIleHisSerGluMetTyrSerLysLeuIleGluThrTyrIleArg 177
QY 442 GATCCCAAGAAAGGAATTTTATTATGATGCAATTTGAAACCATGCGCTTATGTTAAAGAA 501
Db 178 AspGluThrGluArgAsnThrLeuPheAsnAlaValAspGluPheGluPheIleLysLys 197
QY 502 AAAGCAGATGGCGCTTGGATGGATGATGATGATGATGATGATGATGATGATGATGATGATG 561
Db 198 LysAlaAspTrpAlaLeuArgTrpIleSerAspLysLysAlaSerPheAlaGluArgLeu 217
QY 562 GTGGCTTGTCTGTAGAGGATTTTCTCTCAGAGATCTTTGCTGCTATATTTCTGG 621
Db 218 IleAlaPheAlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrp 237
QY 622 CTAAGAAGAGAGGTCTTATGCGCAGGACTCCTTTTCCAAATCACTATCAGCAGAGAT 681
Db 238 LeuLysLysArgGlyLeuMetProGlyLeuThrHisSerAsnGluLeuIleSerArgAsp 257
QY 682 GAAGACTTCATGCTGACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 741
Db 258 GluGlyLeuHisArgAspPheAlaCysLeuLeuTyrSerLysLeuGlnLysLysLeuThr 277
QY 742 GAAGAAGGGTCAGGAGATCATTTGATGCTGCTCAAAATTTGAGCAGGAGTTTAAACA 801
Db 278 GlnGlnArgIleAspIleLysAspAlaValAlaIleGluGlnGluPheLeuThr 297
QY 802 GAAGCTTGGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 861
Db 298 GluAlaLeuProValAspMetIleGlyMetAsnCysArgLeuMetSerGlnTyrIleGlu 317
QY 862 TTTGAGCTGACAGATTTCTGTTGGAATTTCTCAAGTTCTCAAGTTCTTTCAGCAGAAAT 921
Db 318 PheValAlaAspHisLeuLeuValGluLeuGlyCysAspLysLeuTyrLysSerLysAsn 337
QY 922 CCTTTGATTTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGA 981
Db 338 ProPheAspPheMetGluAsnIleSerIleAspGlyLysThrAsnPhePheGluLysArg 357
QY 982 GTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACACAGATAACGCTTTCACCTTG 1041
Db 358 ValSerGluTyrGlnArgProGlyValMetValAsnGluAlaGluArgGlnPheAspLeu 377
QY 1042 GATGAGATTTT 1053
Db 378 GluAlaAspPhe 381
```

RESULT 14

```
US-10-128-714-3245
; Sequence 3245, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
```

```
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3245
; LENGTH: 381
; TYPE: PRF
; ORGANISM: Aspergillus fumigatus
US-10-128-714-3245
```

```
Alignment Scores:
Pred. No.: 3,31e-119 Length: 381
Score: 1192.50 Matches: 229
Percent Similarity: 80.23% Conservative: 51
Best Local Similarity: 65.62% Mismatches: 50
Query Match: 64.60% Indels: 19
DB: 14 Gaps: 4
```

US-10-698-228-12 (1-1053) x US-10-128-714-3245 (1-381)

```
QY 46 GAGAGATCATCTTCAGACACCAAGAAAGTGAATAAGTCAATGAAGAGCCACTCTCTA 105
Db 39 GluAlaSerLysValAlaGlnSerIleLysGluLeuGluAlaAsn---GluProLeuLeu 57
QY 106 AGAAAGAGTTCTCGCGGTTTGTCTCTTCCAAATCCAGTACCCCTGATATTGGAAATG 165
Db 58 GlnGluAsnProHisArgPheValLeuPheProIleLysTyrHisGluIleTrpGlnMet 77
QY 166 TATAAAGCAGCACAGGCTTCTCTCGACACGACAGAGAGGTTGACTTATCAAGAGATCTC 225
Db 78 TyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluIleAspLeuSerLysAspLeu 97
QY 226 CCTCAGTGG---AACAGCTTAAAGCAGATGAGAAAGTACTTTCATCTCTCAGATCTTAGCC 282
Db 98 HisAspTrpAsnAsnArgLeuAsnAspGluArgTyrPheIleSerHisValLeuAla 117
QY 283 TTTTTCAGCCAGTATGGAATTTGTAATGTAATTTGGTGGAGCGCTTTAGTCAGGAG 342
Db 118 PhePheAlaAlaSerAspGlyIleValAsnGluAsnPhe-----AsnGlu 132
QY 343 GTCAGAGTTCACAGGCTCGCTGTTCTATGGCTTTTCAAAATTTCTCATCGAGAATGTTCAAC 402
Db 133 ValGlnValProGluAlaArgCysPheTyrGlyPheGlnIleMetIleGluAsnIleHis 152
QY 403 TCAGAGATGACAGTTTGTGCTAGACACTTACATCAGAGATCCCAAGAAAGGAAATTT 462
Db 153 SerGluThrTyrSerLeuLeuIleAspThrTyrIleLysGluProLysGlnArgPheTyr 172
QY 463 TTATTTAATGCAATTTGAAACCATGCTTATGTAAGAAAGAAAGCAGATTTGGCGCTTGGCA 522
Db 173 LeuPheAspAlaIleAspThrIleProCysIleArgLysLysAlaGlnTrpAlaMetArg 192
QY 523 TGGATAGCAGATAGAAAAATCTACTTTTGGGGAAGAGTGGTGGCTTTTGGCTGCTAGAA 582
Db 193 TrpIleSerAspLysGluSerThrPheAlaGlnArgLeuValAlaPheAlaAlaValGlu 212
QY 583 GGAGTTTCTCTCAGGATCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTTATG 642
Db 213 GlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArgGlyLeuMet 232
QY 643 CAGGAGTCACTTTTCCAAATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTT 702
```

```
|||||
233 ProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisThrAspPhe 252
QY GCTTGCCTGATGTTCCAACTAGTAATAAGCCTTCAGAGAAAGGTCAGGAGATC 762
Db |||||
253 AlaCysLeuLeuPheSerHisLeuAsnHisArgProAspLysLysValIleGluAspIle 272
QY |||||
253 AlaCysLeuLeuPheSerHisLeuAsnHisArgProAspLysLysValIleGluAspIle 272
Db |||||
763 ATTGTTGATGCTGTCACAAATTCAGCAGGAGTTCCTTAAACAGAGCCTTGCCAGTTGCCCTC 822
Db |||||
273 IleValGluAlaValAlaIleGluGlnGluPheLeuThrAspAlaLeuProValAlaLeu 292
QY |||||
823 ATGGAATGAATTGCAATTTGATGAACAGTACATTCAGTTTGTAGTCACAGATTACTT 882
Db |||||
293 LeuGlyMetAsnSerLysLeuMetCysGlnTyrIleGluPheValAlaAspArgLeuLeu 312
QY |||||
893 GTGGAACTTGGATTCTCAAGTTCCTTCAAGTTCCTTCAAGTTCCTTCAAGTTCCTTCAAGTTC 942
Db |||||
313 ValAlaLeuGlyAsnLysLysTyrPheAsnSerProAsnProPheAspPheMetGluSer 332
QY |||||
943 ATTTCTTTAGAGAAAGAAACAAATTTCTTTAGAGAAAGTTCAGAGTATCAGCGCTTTT 1002
Db |||||
333 IleSerLeuAlaGlyLysThrAsnPheGluLysArgValGlyAspTyrGlnLysAla 352
QY |||||
1003 GCAGTTATGGCAGAAACC-----ACAGAT 1026
Db |||||
353 GlyValMetAlaSerThrLysLysAspAlaSerGlnAspThrAlaLysAspAlaAsnAsp 372
QY |||||
1027 AACGTCTTCACCTTCGATGCAGATT 1053
Db |||||
373 GlyGlyLeuCysPheAspGluAspPhe 381
RESULT 15
US-10-032-585-7537
; Sequence 7537, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 7537
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7537
Alignment Scores:
Pred. No.: 3,86e-119 Length: 413
Score: 1192.00 Matches: 225
Percent Similarity: 80.17% Conservative: 58
Best Local Similarity: 63.74% Mismatches: 64
Query Match: 64.57% Indels: 6
DB: 14 Gaps: 3
US-10-698-228-12 (1-1053) x US-10-032-585-7537 (1-413)
QY 13 GAAAGCGCGAAGCGCGCGTGGATCAGGATCAGATGATCTTCAGACACCAACGAA 72
Db |||||
61 GluLysAspSerLeuLysLysHisGlnAspPheLeuAlaLysHisLysValHisArg 80
QY 73 AGTGAATAAAG-----TCAATGAGAGCCACTCTTAAGAAAGAGTTCGCGCGTTT 126
Db |||||
81 HisLysLeuLysGlnLeuGluAlaGluProLeuValGluAsnLysArgTyr 100
QY 127 GTCATCTTCCCAATCCAGTACCTGATATTTGGAAATGTATAAAGCAGGACAGGCTTC 186
Db |||||
101 ValMetPheProIleArgTyrHisGluIleThrAsnPheTyrLysLysAlaGluAlaSer 120
```

```
187 TTCTGCACGACGAGAGGTTGACTTATCAAGGATCTCCCTCACTGG- - -AACAGCTT 243
Db |||||
121 PheTrpThrAlaGluGluIleAspLeuSerLysAspLeuAspTrpAsnAsnLysLeu 140
QY |||||
244 AAACGACATCAGAAAGTACTTCTCTCATCATCTTACGCTTTTTCAGCAGGATGATGGA 303
Db |||||
141 AsnGluAsnGluArgTyrPheIleSerArgValLeuAlaPhePheAlaSerAspGly 160
QY |||||
304 ATTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGAGGTCAGAGGCTCGC 363
Db |||||
161 IleValGlyGluAsnLeuIleGluAsnPheSerThrGluValGlnLeuProGluAlaLys 180
QY |||||
364 TGTTCCTATCGCTTCAAAATTCATCATCGAATCTTCACTCAGAGATCTCAGATTGTGTC 423
Db |||||
181 SerPheTyrGlyPheGlnIleMetMetGluAsnIleHisSerGluThrTyrSerLeuLeu 200
QY |||||
424 ATAGACACTTACATCAGAGATCCCAAGAAAGGAATTTTATTATTAATGCAATTGAAACC 483
Db |||||
201 IleGluThrTyrIleLysAspProGlnGluAlaAspTyrLeuPheAsnAlaIleAlaAsn 220
QY |||||
484 ATGCCCTATGTTAAGAAAGACGATTTGGCCTTGCATGATGATGATGATGATGATGATGAT 543
Db |||||
221 IleProCysIleGlnLysLysAlaAspTrpAlaIleLysTrpIleGlnAspAspGluAla 240
QY |||||
544 ACTTTTGGGAAAGAGTGGTGGCTTCTGCTCTAGAGAGGATTTCTTCTCAGGATCT 603
Db |||||
241 LeuTyrGlyGluArgLeuValAlaPheAlaAlaValGluGlyIlePhePheSerGlySer 260
QY |||||
604 TTTGCTGCTATATTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACATTTTCCAAAT 663
Db |||||
261 PheAlaSerIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSerAsn 280
QY |||||
664 GAACCTCATCAGCAGATGAAAGACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 723
Db |||||
281 GluLeuIleCysArgAspGluGlyLeuHisThrAspPheAlaCysLeuLeuPheAlaHis 300
QY |||||
724 TTAGTAAATAAGCTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAAATT 783
Db |||||
301 LeuGlnAsnArgProSerProGluIleValGluArgIleIleThrGluAlaValAspVal 320
QY |||||
784 GACGAGAGTTCCTTAAACAGAGCCTTCCAGCTTCCCTCATTTGGAATGAAATTCGATTTTG 843
Db |||||
321 GluLysGluTyrPheThrAspAlaLeuProValSerLeuLeuGlyMetAsnCysAsnLeu 340
QY |||||
844 ATGAAAACAGTACATTCAGTTGTTAGCTGACAGATTCATTTGTTGGAACCTTCTCAAG 903
Db |||||
341 MetCysGlnTyrIleGluPheValAlaAspArgLeuLeuLeuAlaLeuGlyAsnLysLys 360
QY |||||
904 GTTTTTCAGCAGCAAAATCCTTTTATTTATGGAACCAATTTCTTTAGAGGAAACCA 963
Db |||||
361 ValTyrAsnValThrAsnProPheAspPheMetGluAsnIleSerLeuAlaGlyLysThr 380
QY |||||
964 AATTTCTTTCAGAAACGAGTTTCAGAGTATCAGGCTTTTTCAGGATTCATGCGCAACCA 1023
Db |||||
381 AsnPhePheGluLysArgValSerAspTyrGlnLysAlaGlyValMetAlaLysThrGlu 400
QY 1024 GAT-----AACGTCTTTCACCTTCGATGTCAGATT 1053
Db |||||
401 AsnLysGluAlaAspAlaPheThrPheAspGluAspPhe 413
```

Search completed: October 30, 2005, 06:55:20
Job time : 80.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 30, 2005, 06:24:21 ; Search time 173.25 Seconds
(without alignments)
14198.102 Million cell updates/sec

Title: US-10-698-228-12

Perfect score: 1846

Sequence: 1 atggcgaccggaagcc.....tcaccttgatgcagatttt 1053

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6959266 seqs, 1168006243 residues

Total number of hits satisfying chosen parameters: 13918532

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame_n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10698228/runat_26102005_100609_4045/app_query.fasta_1.2446
-DB=Pending Patents AA Main -QFMT=fastan -SUFFIX=n2p.rapm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEARSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10698228.cgn2_1_523 @runat_26102005_100609_4045
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Pending Patents AA Main:
1: /cgn2_6/ptodata/1/paa/PCTUS COMB.pap.*
2: /cgn2_6/ptodata/1/paa/US06 COMB.pap.*
3: /cgn2_6/ptodata/1/paa/US07 COMB.pap.*
4: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
5: /cgn2_6/ptodata/1/paa/US081 COMB.pap.*
6: /cgn2_6/ptodata/1/paa/US082 COMB.pap.*
7: /cgn2_6/ptodata/1/paa/US083 COMB.pap.*
8: /cgn2_6/ptodata/1/paa/US084 COMB.pap.*
9: /cgn2_6/ptodata/1/paa/US085 COMB.pap.*
10: /cgn2_6/ptodata/1/paa/US086 COMB.pap.*
11: /cgn2_6/ptodata/1/paa/US087 COMB.pap.*
12: /cgn2_6/ptodata/1/paa/US088 COMB.pap.*
13: /cgn2_6/ptodata/1/paa/US089 COMB.pap.*
14: /cgn2_6/ptodata/1/paa/US090 COMB.pap.*
15: /cgn2_6/ptodata/1/paa/US091 COMB.pap.*
16: /cgn2_6/ptodata/1/paa/US092 COMB.pap.*
17: /cgn2_6/ptodata/1/paa/US093 COMB.pap.*
18: /cgn2_6/ptodata/1/paa/US094 COMB.pap.*
19: /cgn2_6/ptodata/1/paa/US095 COMB.pap.*
20: /cgn2_6/ptodata/1/paa/US096 COMB.pap.*
21: /cgn2_6/ptodata/1/paa/US097A COMB.pap.*
22: /cgn2_6/ptodata/1/paa/US097B COMB.pap.*
23: /cgn2_6/ptodata/1/paa/US098 COMB.pap.*
24: /cgn2_6/ptodata/1/paa/US099A COMB.pap.*
25: /cgn2_6/ptodata/1/paa/US099B COMB.pap.*
26: /cgn2_6/ptodata/1/paa/US100 COMB.pap.*
27: /cgn2_6/ptodata/1/paa/US101 COMB.pap.*
28: /cgn2_6/ptodata/1/paa/US102 COMB.pap.*

29: /cgn2_6/ptodata/1/paa/US103 COMB.pap.*
30: /cgn2_6/ptodata/1/paa/US104 COMB.pap.*
31: /cgn2_6/ptodata/1/paa/US105 COMB.pap.*
32: /cgn2_6/ptodata/1/paa/US106 COMB.pap.*
33: /cgn2_6/ptodata/1/paa/US107 COMB.pap.*
34: /cgn2_6/ptodata/1/paa/US108 COMB.pap.*
35: /cgn2_6/ptodata/1/paa/US109 COMB.pap.*
36: /cgn2_6/ptodata/1/paa/US110 COMB.pap.*
37: /cgn2_6/ptodata/1/paa/US60 COMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1821	98.6	351	22	US-09-787-491B-11
2	1821	98.6	351	22	US-09-791-537-77474
3	1821	98.6	351	27	US-10-170-205E-19045
4	1821	98.6	351	32	US-10-698-228-1
5	1821	98.6	351	35	US-10-990-328-7547
6	1821	98.6	351	37	US-60-128-660-7
7	1821	98.6	351	37	US-60-449-629-94
8	1821	98.6	351	37	US-60-505-218-322
9	1821	98.6	366	22	US-09-791-537-27671
10	1817	98.4	351	20	US-09-629-469A-12811
11	1817	98.4	351	22	US-09-791-537-112185
12	1817	98.4	351	35	US-10-917-503-12811
13	1454	78.8	389	1	PCT-US02-10824-143
14	1454	78.8	389	1	PCT-US03-17409-171
15	1454	78.8	389	10	US-08-663-617-33
16	1454	78.8	389	12	US-08-837-599-33
17	1454	78.8	389	27	US-09-791-537-88179
18	1454	78.8	389	27	US-10-170-205E-8018
19	1454	78.8	389	30	US-10-408-765-559
20	1454	78.8	389	30	US-10-408-765A-559
21	1454	78.8	389	32	US-10-698-228-5
22	1454	78.8	389	33	US-10-733-878-457
23	1454	78.8	389	33	US-10-756-149-4899
24	1454	78.8	389	35	US-10-940-774-6016
25	1454	78.8	389	37	US-60-389-987-6016
26	1454	78.8	389	37	US-60-412-418-559
27	1454	78.8	389	37	US-60-452-680-20616
28	1454	78.8	389	37	US-60-568-073-984
29	1454	78.8	413	1	PCT-US00-05882-1347
30	1454	78.8	413	24	US-09-925-301-1347
31	1454	78.8	449	35	US-10-990-328-12349
32	1454	78.8	449	35	US-10-990-328-12350
33	1454	78.8	453	1	PCT-US01-04926A-686
34	1454	78.8	453	1	PCT-US01-08631-45646
35	1454	78.8	453	28	US-10-220-335-686
36	1454	78.8	453	28	US-10-290-752-686
37	1454	78.8	453	35	US-10-940-774-7896
38	1443	78.2	390	10	US-08-663-617-35
39	1443	78.2	390	12	US-08-837-599-35
40	1443	78.2	390	22	US-09-791-537-117257
41	1425.5	77.2	386	22	US-09-791-537-114288
42	1425.5	77.2	386	30	US-10-403-571-76
43	1425.5	77.2	386	30	US-10-405-887-76
44	1421	77.0	386	22	US-09-791-537-36076
45	1349.5	73.1	399	22	US-09-791-537-1129

ALIGNMENTS

RESULT 1

US-09-787-491B-11

; Sequence 11, Application US/09787491B

; GENERAL INFORMATION:

; APPLICANT: INCYTE GENOMICS, INC.; TANG, Y. Tom;

; APPLICANT: CORLEY, Neil C.; GUEGLER, Karl J.;
 ; APPLICANT: GORONE, Gina A.; PATTERSON, Chandra;
 ; APPLICANT: HILLMAN, Jennifer L.; BAUGHN, Mariah R.;
 ; APPLICANT: LAL, Preeti; AZIMZAI, Yalda;
 ; APPLICANT: YUE, Henry; YANG, Junming
 ; TITLE OF INVENTION: RNA-ASSOCIATED PROTEINS
 ; FILE REFERENCE: PF-0600 USN
 ; CURRENT APPLICATION NUMBER: US/09/787,491B
 ; CURRENT FILING DATE: 2002-11-04
 ; PRIOR APPLICATION NUMBER: PCT/US99/21688
 ; PRIOR FILING DATE: 1999-09-17
 ; PRIOR APPLICATION NUMBER: US 60/128,660
 ; PRIOR FILING DATE: 1999-04-08
 ; PRIOR APPLICATION NUMBER: US 60/069,391
 ; PRIOR FILING DATE: 1998-11-04
 ; PRIOR APPLICATION NUMBER: US 60/183,025
 ; PRIOR FILING DATE: 1998-09-22
 ; PRIOR APPLICATION NUMBER: US 60/155,246
 ; PRIOR FILING DATE: 1998-09-17
 ; PRIOR APPLICATION NUMBER: US 09/158,720
 ; PRIOR FILING DATE: 1998-09-17
 ; NUMBER OF SEQ ID NOS: 38
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 11
 ; LENGTH: 351
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No: 2604449CD1
 US-09-787-491B-11

Alignment Scores:
 Pred. No.: 2,58e-186 Length: 351
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.65% Indels: 0
 DB: 22 Gaps: 0

US-10-698-228-12 (1-1053) x US-09-787-491B-11 (1-351)

QY 1 ATGGCGACCCGGAAGGCGGAGCGGCTGGATCAGGATGAGATCATCTTCA 60
 DB 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
 QY 61 GACACCAACGAAGTGAATAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
 DB 21 AspThrAsnGluSerGluLeuLeuSerAsnGluProLeuLeuArgLysSerSerArg 40
 QY 121 CGGTTTGTCATCTTCCATCCAGTACCTGATATTTGGAAATGTATAACAGGCACAG 180
 DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
 QY 181 GCTTCTCTCTGACAGCAGCAAGAGGTGACTTATCAAGATCTCCCTCACTGGAACAAG 240
 DB 61 AlaSerPheTrpThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 QY 241 CTTAAAGCAGATGAGAAGTACTTCTCTCTCATCTTACATCTTACGCTTTTTCAGCCAGTGAT 300
 DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisLeuAlaPhePheAlaAlaSerAsp 100
 QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
 DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnProGluAla 120
 QY 361 CGCTGTTCTATGGCTTTCAATTTCTCATCGAGAATGTTCACTCAGAGATGACAGTTTG 420
 DB 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGAA 480
 DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160

QY 481 ACCATGCCCTATCTTAAGAAAAAAGCAGATTGGGCCCTTGCATGATAGCAGATAGAAA 540
 DB 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 QY 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGTTTCTTCTCAGA 600
 DB 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 QY 601 TCTTTTGTCTATATCTCGCTAAAGAGAGGCTTATGCCAGGACTCATTGCTTTTCC 560
 DB 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 QY 661 AATGAACCTCATCAGCAGATGAAGGACTTCACTGTGCTGCTTGTGCTGCTGATGTTCCAA 720
 DB 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 QY 721 TACTTAGTAATAAGCCTTCAGAAAGAGGTGAGGAGATCATTTGTTGATGCTCTCAAA 780
 DB 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 QY 781 ATTGAGCAGAGTTTTAAACAGAGCCTTGCAGTGGCTCATTTGGATGAATGAATTCATT 840
 DB 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 QY 841 TTGATGAACAGTACATTTGAGTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
 DB 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 QY 901 AAGGTTTTTCAGGCAGAAAAATCCTTTTGTATTTATGAAAAACATTTCTTTAGAGGAAA 960
 DB 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 QY 961 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGAGTTTATGCGCAGAAACC 1020
 DB 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 QY 1021 ACAGATAACCTCTTCACCTTGGATGACAGATTTT 1053
 DB 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 2

US-09-791-537-77474

; Sequence 77474, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Bionomix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBER

; TITLE OF INVENTION: METHODS OF USE THEREOF

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 77474

; LENGTH: 351

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-791-537-77474

Alignment Scores:

Pred. No.: 2,58e-186 Length: 351
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.65% Indels: 0
 DB: 22 Gaps: 0

US-10-698-228-12 (1-1053) x US-09-791-537-77474 (1-351)

QY 1 ATGGCGACCCGGAAGGCGGAGCGGCTGGATCAGGATGAGATCATCTTCA 60
 DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160

Db 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSerSer 20
QY 61 GACACCAACGAAGTGAATAAGTCAATGAAGAGCCACTCTAAGAAGAGTTCTCGC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
QY 121 CGTTTGTTCATCTTCCAAATCCAGTACCTGATATTTGGAAAATGATATAACAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTTCGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCAGTGAACAAG 240
Db 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAAGCAGATGAGAGTACTTCTCTCATCTTACATCTTACGCTTTTTCAGCCAGTAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTCTATGGCTTTCAAAATTCATCGAGAAATGTTCACTCAGAGATGACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 160
QY 481 ACCATGCCCTATGTTAAAGAAAAGCAGATTCGGGCTTCGGATGGATAGCAGATAGAAAA 540
Db 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAAGAAGGATTTTCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGluValPhePheSerGly 200
QY 601 TCTTTTGTCTGCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCATTTTTC 660
Db 201 SerPheAlaAlaIlePheThrLysLysAlaValGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACATCATCAGCAGATGAAGGACTTCACTGTGACTTTGTGCTTGTGCTGCTGCTCA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240

Db 1021 ACAGATAACGCTTTCACCTTGCATCGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 3
US-10-170-205E-19045
; Sequence 19045, Application US/10170205E

; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19045
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-19045

Alignment Scores: 2.58e-186 Length: 351
Pred. No.: 1821.00 Matches: 351
Score: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 27 Gaps: 0

US-10-698-228-12 (1-1053) x US-10-170-205E-19045 (1-351)

QY 1 ATCGGCGACCCCGAAAGCGCGGAGCGCGGGCTGTGATCAGGATGAGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSerSer 20
QY 61 GACACCAACGAAGTGAATAAGTCAATGAAGAGCCACTCTAAGAAGAGTTCTCGC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
QY 121 CGTTTGTTCATCTTCCAAATTCACATCCAGTACCTTCAATGATTTTGGAAAATGATATAACAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTTCGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCAGTGAACAAG 240
Db 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAAGCAGATCAGAGTACTTCTCATCTCTCACATCTTAGCGCTTTTTCAGCCAGTAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTCTATGGCTTTCAAAATTCATCAGAGATGTTCACTCAGAGATGACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTTAATGCAATTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 160
QY 481 ACCATGCCCTATGTTAAAGAAAAGCAGATTCGGGCTTCGGATGGATAGCAGATAGAAAA 540
Db 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAAGAAGGATTTTCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGluValPhePheSerGly 200
QY 601 TCTTTTGTCTGCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCATTTTTC 660
Db 201 SerPheAlaAlaIlePheThrLysLysAlaValGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACATCATCAGCAGATGAAGGACTTCACTGTGACTTTGTGCTTGTGCTGCTGCTCA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240

```
QY 721 TACTTAGTAATAAGCTTTCAGAGAAAGGTCAGGAGATCATTTGATGCTGTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluLeuIleValAspAlaValLys 260
QY 781 ATTGAGCAGGAGTTTAAACAGAGCTTGCAGTTGGCTCATTTGGAATGAATTCGATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACACACTACATTCAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTCCTCA 900
Db 281 LeuMetLysGlnIleGluPheValAlaAspArgLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGCGAGAAATCCTTTGATTTATGGAAACATTTCTTTAGAGCAAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluLys 320
QY 961 ACAATTTCTTGAGAACAGGTTTCAGAGTATCAGCGTTTTCAGATTGCGAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTCTTCACCTTGGATGCAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 4

```
US-10-698-228-1
; Sequence 1, Application US/10698228
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 1
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-1
```

```
Alignment Scores:
Pred. No.: 2,58e-186 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
Db: 32 Gaps: 0
```

US-10-698-228-12 (1-1053) x US-10-698-228-1 (1-351)

```
QY 1 ATGGCGCCCGGAAGGCGGAGCGCGGCTGGATCAGATCAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAGTGAATAAGTCAAAATGAAGCCACTCTTAAGAAGAGTTCTCGC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGTTTGTTCATCTTCCAAATCCAGTACCTGTATTTGGAAATGATTAACAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTTCTGACACAGCAGAGAGGTTGACTTATCAAGGATCTCCTCACTGGAACAAG 240
```

```
Db 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAAGCAGATCAGAGTACTTCATCTCTCACATCTTACGCTTTTTCGACCCAGTCAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisLeuAlaPhePheAlaSerAsp 100
QY 301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATCGCTTTCAAATTTCTCATCGAATGTTCTCACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGATCCCAAGAAAGGAATTTTATTAAATGCAATTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCCCTTGGCATGGATGAGATAGAAAA 540
Db 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTCTGCTGTAGAGAGGATTTTCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGGCTTATGCCAGAGACTCATTCTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACTCATCAGCAGAGATGAAGACTTCACGTGCTGCTTGTGCTGCTGCTGCTGCT 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATAAGCTTTCAGAGAAAGGTCAGGAGATCATTTGATGCTGTCTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGAGCAGGAGTTTAAACAGAGCTTGCAGTTGGCTCATTTGGAATGAATTCGATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACACACTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAATTCGATTCCTCA 900
Db 281 LeuMetLysGlnIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGCGAGAAATCCTTTGATTTATGGAAACATTTCTTTAGAGCAAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluLys 320
QY 961 ACAATTTCTTGAGAACAGGTTTCAGATTCAGCGTTTTCAGATTGCGAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTCTTCACCTTGGATGCAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 5

```
US-10-990-328-7547
; Sequence 7547, Application US/10990328
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; FILE REFERENCE: CLO01495
; CURRENT APPLICATION NUMBER: US/10/990,328
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7547
```



```
Db      121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY      421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATCAATTTGAA 480
Db      141 LeuIleAspThrTyrIleArgaspProLysLysArggluPheLeuPheAsnAlaIleGlu 160
QY      481 ACCATGCCCTATCTTAAAGAAAAAGCAGATTGGCCCTTCGATGGATAGCAGATAGAAAA 540
Db      161 ThrMetProTyrValLysLysLysAlaAspTAlaLeuArgTrpIleAlaaspArgLys 180
QY      541 TCTACTTTTGGGGAAGAGTGTGGCTTTGCTGCTGTGTAGAGAGATTTTCTCTCAGGA 600
Db      181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY      601 TCTTTTGTCTATATTCTGGCTTAAAGACAGAGGCTTTATGCCAGACTCTCTTTTCTCC 660
Db      201 SerPheAlaAlaIlePheTrpLeuLysLysArgLysLeuMetProGlyLeuThrPheSer 220
QY      661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGGCTGTGTTCCAA 720
Db      221 AsnGluLeuIleSerArgaspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY      721 TACTTAGTAATAAAGCTTTTACAGAAAGGCTCAGGAGATCATTTGTGCAACTTTGATTTCTCA 900
Db      241 TyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAlaValLys 260
QY      781 ATTGAGCAGGAGTCTTACAGAGAGCTTCCAGAGCTTCCAGTGTGCTCATTTGGAATGTCATT 840
Db      261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY      841 TTGATGAACAGCTCTTCACTTGGATGGATGAGATTTT 1053
Db      341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 7

```
US-60-449-629-94
; Sequence 94, Application US/60449629
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001449
; CURRENT FILING DATE: 2003-02-04
; NUMBER OF SEQ ID NOS: 2712
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 94
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-449-629-94
```

```
Alignment Scores:
Pred. No.: 2,58e-186 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 37 Gaps: 0
```

```
US-10-698-228-12 (1-1053) x US-60-449-629-94 (1-351)
QY      1 ATGGCGACCCGAAAGCGCGGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db      1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
QY      61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db      21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
QY      121 CGGTTTGTCTATCTTCCAATCCAGTACCTCGATATTTGGAAAATGTATAAACAGGCACAG 180
Db      41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY      181 GCCTTCTTCGGACAGCAGAGAGGTTGACTTATCAAGAGGATCTCCCTCACTGGAAACAAG 240
Db      61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY      241 CTTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCGACGCCAGTGAT 300
Db      81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY      301 GGAATTTGTAATAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGCT 360
Db      101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY      361 CGCTGTTTCTATCGCTTCAAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db      121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY      421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db      141 LeuIleAspThrTyrIleArgaspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY      481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCCCTTCGATGGATGAGATGAGATAAAAA 540
Db      161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY      541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTCTCTAGAGAGGATTTCTCTCAGGA 600
Db      181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY      601 TCTTTTGTCTATATTCTGGCTTAAAGAAAGAGAGGTCTTATGCCAGGAGCTCACTTTTTC 660
Db      201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY      661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGGCTGTGTTCCAA 720
Db      221 AsnGluLeuIleSerArgaspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY      721 TACTTAGTAATAAAGCTTTTACAGAAAGGCTCAGGAGATCATTTGTGATGCTCTCAAA 780
Db      241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY      781 ATTGAGCAGGAGTCTTAAACAGAAAGCCTTGGCAGTGGCCCTCATTTGGAATCAATTCATT 840
Db      261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY      841 TTGATGAACAGTACATGATGAGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATTTCTCA 900
Db      281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY      901 AAGTTTTTTCAGCAGAAAAATCTTTTGATTTTATGAAAAACATTTCTTTAGAGGAAAA 960
Db      301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluLys 320
QY      961 ACAAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTGCAGTTATGGCAGAAACC 1020
Db      321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY      1021 ACAGATAACGCTCTTCACTTGGATGGATGAGATTTT 1053
Db      341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```


RESULT 8

US-60-505-218-322
; Sequence 322, Application US/60505218
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001492
; CURRENT FILING DATE: 2003-09-24
; NUMBER OF SEQ ID NOS: 22507
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 322
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-505-218-322

Alignment Scores:
Pred. No.: 2,58e-186 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 37 Gaps: 0

US-10-698-228-12 (1-1053) x US-60-505-218-322 (1-351)

QY	1	ATGGGGGACCCGAAAGCCGGAAGCCGGCGCGGTGATCAGGATGAGAGATCATCTTCA	60
DB	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer	20
QY	61	GACACCAACGAAGTGAATAAAGTCAAAATCAAGACGACCCTCTTAAGAAAGAGTTCTCG	120
DB	21	AspThrAsnGluSerGluIleYserAsnGlnGluProLeuLeuArgLysSerSerArg	40
QY	121	CGGTGTTGCATCTTTCCAAATCCAGTACCCTCATATTTGGAAAATGTATAACAGGCACAG	180
DB	41	ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln	60
QY	181	GCTTCTTCTGACAGCAGAGAGAGTTGACTTATCAAAAGATCTCCCTCACTCGAAACAG	240
DB	61	AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys	80
QY	241	CTTAAGCAGATGACAAGTACTTCTATCTCCACATCTTAGCCCTTTTTCGAGCCAGTGAT	300
DB	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp	100
QY	301	GGAAATGTAAATGAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT	360
DB	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla	120
QY	361	CGCTGTTCTATGGCTTTTCAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG	420
DB	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTATGCAATTGAA	480
DB	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160
QY	481	ACCATGCCCTATGTTAAGAAAAAGACAGATTTGGGCCTTGCATGGATAGCAGATAGAAA	540
DB	161	ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys	180
QY	541	TCTACTTTTGGGAAAGAGTGGTGGCCTTTTGCTGCTAGAGGAGGTTTCTTCTCAGGA	600
DB	181	SerThrPheGlyGluArgValAlaAlaPheAlaAlaValGluGlyValPhePheSerGly	200
QY	601	TCTTTTGCTGCTATATCTTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCATTCTTTC	660
DB	201	SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer	220

Qy	661	AATGAAC	TCA	TACAGCAGAGATGAA	GACTTC	CTACGTG	CACTTTGC	TTCCTCGCTGATGTTCCAA	720
Db	221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln							240
Qy	721	TACTTACTAATAAGCCTTCAGAGAAGAAAGGGTCAGGGAGATCAT							780
Db	241	TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys							260
Qy	781	ATTGAGCAGGAGTTTTTAACAAGAACGCTTGCCACAGTTGGCCCTCAT							840
Db	261	IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle							280
Qy	841	TTGATGAAAACAGTACATGATTTGTAGCTGACAGATTA							900
Db	281	LeuMetLysGlnTrpIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer							300
Qy	901	AAGGTTTTTTCAGGCAGAAATCCCTTTTGATTTTATGAAAAACATTTCTTTAGAGGAAAA							960
Db	301	LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys							320
Qy	961	ACAAATTTCTTTCAGAAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC							1020
Db	321	ThrAsnPhePheGluLysArgValSerGluTrpGlnArgPheAlaValMetAlaGluThr							340
Qy	1021	ACAGATAACGCTCTTCACCTTGGATGCAGATTTT							1053
Db	341	ThrAspAsnValPheThrLeuAspAlaAspPhe							351
RESULT 9									
US-09-791-537-27671									
; Sequence 27671, Application US/09791537									
; GENERAL INFORMATION:									
; APPLICANT: Bionomix, Inc.									
; APPLICANT: Debe, Derek									
; APPLICANT: Danzer, Joseph									
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAM									
; FILE REFERENCE: 261/210									
; CURRENT APPLICATION NUMBER: US/09/791,537									
; CURRENT FILING DATE: 2001-02-22									
; NUMBER OF SEQ ID NOS: 153055									
; SOFTWARE: PatentIn version 3.0									
; SEQ ID NO 27671									
; LENGTH: 366									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-09-791-537-27671									
Alignment Scores:									
Pred. No.:	2,62e-186	Length:	366						
Score:	1821.00	Matches:	351						
Percent Similarity:	100.00%	Conservative:	0						
Best Local Similarity:	100.00%	Mismatches:	0						
Query Match:	98.65%	Indels:	0						
DB:	22	Gaps:	0						
US-10-698-228-12 (1-1053) x US-09-791-537-27671 (1-366)									
Qy	1	ATGGGGACCCGAAAGGCCGGAAGCGCGGGCTGGATCAGATCAGATGATCATCTTCA							60
Db	16	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer							35
Qy	61	GACACCAAGAAAGTCAAATGAAGTCAAATGAAGACCACTCTTAAGAAGAGTCTTCGC							120
Db	36	AspThrAsnGlnSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg							55
Qy	121	CGGTTTGTCACTTTTCCAATCCAGTACCCCTGATATTGGAAAAATGATATAACAGGCACAG							180
Db	56	ArgPheValIlePheProIleGlnTrpProAspIleTrpLysMetTrpLysGlnAlaGln							75
Qy	181	GCTTCCTTCGACACAGAAAGGTTGACTTATCAAGAGTCTCCCTCACTCGAAACAAG							240
Db	76	AlaserPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHISTDAsnLys							95

```
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCTCATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db 96 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaSerAsp 115
QY 301 GGAATTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTCCAGAGCT 360
Db 116 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 135
QY 361 CGCTGTTCTATGGCTTTCCTCAATCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTG 420
Db 136 ArgCysPheTyrGlyPheGlnIleuIleGluAsnValHisSerGluMetTyrSerLeu 155
QY 421 CTGATAGACACTTATACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 480
Db 156 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 175
QY 481 ACCATGCCCTATGTTAAAGAAAAAGCAGATTTGGCCCTTCGATGATAGATAGAAAA 540
Db 176 ThrMetProTyrValLysLysLysAlaAspTTPAlaLeuArgTTPileAlaAspArgLys 195
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
Db 196 SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 215
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCCAGACTCACTTTTTC 660
Db 216 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 235
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGTCTTGTCTGATGTTCCAA 720
Db 236 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 255
QY 721 TACTTAGTAAATAGCTTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 256 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAlaValLys 275
QY 781 ATTGAGCAGAGTATTTTACAGAGCTTTCAGAGTGGCTTCCAGTGGCTCATTCGAAATGCAATT 840
Db 276 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 295
QY 841 TTGATGAAACAGTACATGAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTTCTCA 900
Db 296 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 315
QY 901 AAGTTTTTCAGCAGAGAAATCCTTTGATTTTATGGAACCAATTTCTTTAGAGGAAAA 960
Db 316 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 335
QY 961 ACAAAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
Db 336 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 355
QY 1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053
Db 356 ThrAspAsnValPheThrLeuAspAlaAspPhe 366
```

RESULT 10

```
US-09-629-469A-12811
; Sequence 12811, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
```

```
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/09/629,469A
; FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12811
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-629-469A-12811
```

```
Alignment Scores:
Pred. No.: 6 97e-186 Length: 351
Score: 1817.00 Matches: 350
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.72% Mismatches: 0
Query Match: 98.43% Indels: 0
DB: 20 Gaps: 0
```

US-10-698-228-12 (1-1053) x US-09-629-469A-12811 (1-351)

```
QY 1 ATGGCGCACCGGAAAGCCGCGAGCGCGGTGATCATGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACAAAGTGAATAAAGTCAAAATGAAGAGCCACTCCTAAGAAAGTTCTGCG 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGGTTTGTCTCTTCCCAATCCAGTACCTGATATTTGGAAAAATCTATAAACAGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTTCGACAGCAGAGAGTTGACTTATCAAGAGATCTCCCTCCTCAGTGAACAG 240
Db 61 AlaSerPheTyrThrAlaGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCAGATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
QY 301 GGAATTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTCCAGAGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTCTATGGCTTTCCTCAATCTCATCAGAAATGTTCTACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTATCAGATGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAAGAAAAAGCAGATTTGGCCCTTCGATGATAGATAGAAAA 540
Db 161 ThrMetProTyrValLysLysAlaAspTTPAlaLeuArgTTPileAlaAspArgLys 180
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
```

Db 181 SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGGCTATATCTGGCTAAAGAGAGAGGCTTATGCGAGGACTCCTTTTCC 660
Db 201 SerPheAlaAlaIlePheIlePheIlePheIlePheIlePheIlePheIlePheIle 220
QY 661 ATGAACCTCATGACGAGATGAAGACTTCACTGTGACTTTCCTGCTGCTGCTGCTGCT 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATAAGCTTCAGAGAAAGGCTCAGGAGATCATTTGCTGCTGCTGCTGCT 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAlaValLys 260
QY 781 ATTGACGAGAGTCTTAAAGAGGCTTGCAGGCTTGCAGGCTTGCAGGCTTGCAGGCT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyValAsnCysIle 280
QY 841 TTGATGAACACGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AGGTTTTCAGGACGAGAAATCTTTGATTTTATGGAACCAATTTCTTTAGAGGAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTTCAGAAACGCTTTCAGAGTATCAGGCTTTCAGGCTTTCAGGCT 1020
Db 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACCTCTTACCTTGGATGCGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 11

US-09-791-537-112185
; Sequence 112185, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 112185
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-112185

Alignment Scores:

Pred. No.:	6,97e-186	Length:	351
Score:	1817.00	Matches:	350
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.72%	Mismatches:	0
Query Match:	98.43%	Indels:	0
DB:	22	Gaps:	0

US-10-698-228-12 (1-1053) x US-09-791-537-112185 (1-351)

QY 1 ATGGGCGACCCGGAAGCGGAGCGGCTGAGTACAGGATGAGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAGTGAATAAGTCAATGAAGACCACTCCTAAGAAGAGTCTCGC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGTTTGTGTCATCTTCCATCCAGTACCTGTATATTTGGAAATGTATAAACAGGCAC 180

Db 41 ArgPheValIlePheProIleGlnTyrProAspIleIleTyrLysGlnAlaGln 60
QY 181 GCCTTCTCTGGACAGAGAGAGGCTTCACTATCAAGAGATCTCCCTCAGTGGACAAG 240
Db 61 AlaSerPheThrPheAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAACGAGATGAGAGTACTTCACTCTCACAATCTTAGCCTTTTGTGACGAGTGTAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
QY 301 GGAATCTTAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTCAGGTTCCAGAGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTCTATGCTGCTTCAAAATCTCATCTGAGTGTCTCATGAGATGTACAGATTGTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAAATGCAATTGAA 480
Db 141 LeuIleAspThrTyrIleArgPheLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCTTATGTTAAAGAAAAACAGATTTGGGCTTTCGATGATGATGATGATGATGAT 540
Db 161 ThrMetProTyrValLysLysLysLysLysLysLysLysLysLysLysLysLysLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGGCTATATTTCTGGCTAAAGAGAGGCTTTCATGCGAGGACTCCTTTTTC 660
Db 201 SerPheAlaAlaIlePheThrLeuLysLysLysLysLysLysLysLysLysLysLys 220
QY 661 AATGAACCTCATCAGCAGATGAAGGACTTCACTGTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATAAGCTTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGCTCA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAlaValLys 260
QY 781 ATTGACGAGAGTCTTAAAGAGGCTTGCAGGCTTGCAGGCTTGCAGGCTTGCAGGCT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyValAsnCysIle 280
QY 841 TTGATGAACACGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGTGTTCCTCAGCAGAAAAATCCTTTTGTATGTAATGTAATGTAATGTAATGTAAT 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTCTTTCAGAAACGAGTTCAGATTCAGGCTTTCAGGCTTTCAGGCTTTCAGGCT 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTTTCACCTTGGATGCGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 12

US-10-917-503-12811
; Sequence 12811, Application US/10917503
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI

APPLICANT: ISHII, SHIZUKO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: WAKAMATSU, AI
APPLICANT: NAGAI, KEIICHI
APPLICANT: OTSUKI, TETSUJI
TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
FILE REFERENCE: 084335/0123
CURRENT APPLICATION NUMBER: US/10/917,503
CURRENT FILING DATE: 2004-08-13
PRIORITY FILING DATE: 2000-07-28
PRIORITY FILING DATE: 2000-07-28
PRIORITY FILING DATE: 1999-07-29
PRIORITY FILING DATE: 1999-07-29
PRIORITY FILING DATE: 1999-08-27
PRIORITY FILING DATE: 2000-01-11
PRIORITY FILING DATE: 2000-01-11
PRIORITY FILING DATE: 2000-05-02
PRIORITY FILING DATE: 2000-05-02
PRIORITY FILING DATE: 2000-06-09
PRIORITY FILING DATE: 2000-06-09
PRIORITY FILING DATE: 1999-10-18
PRIORITY FILING DATE: 2000-02-17
NUMBER OF SEQ ID NOS: 19025
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12811
LENGTH: 351
TYPE: PRT
ORGANISM: Homo sapiens
US-10-917-503-12811

Alignment Scores:
Pred. No.: 6,97e-186 Length: 351
Score: 1817.00 Matches: 350
Percent Similarity: 100.00% Conservativity: 1
Best Local Similarity: 99.72% Mismatches: 0
Query Match: 98.43% Indels: 0
DB: 35 Gaps: 0

US-10-698-228-12 (1-1053) x US-10-917-503-12811 (1-351)

QY 1 ATGGCGCCGCGGAAAGCGCGGAGCGCGGCTGGATCAGATCAGATCATCTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaLaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAAGTCAATGAAATGAAGAGCCACTCTTAAGAGAGAGTCTTCGC 120
DB 21 AspThrAsnGluSerGluLeuLeuSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGTTTGTTCATCTTCCATCCAGTACCTCGATATTTGGAAATGATTAACAGGCAAG 180
DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTTGACAGACAGAGGTTGACTTATCAAGATCTCCCTCACTGGACACAG 240
DB 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAAGCAGATGAGAGTACTTCTCATCTTACATCTTACATCTTTCGAGCCAGTAT 300
DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGCAGGTTCCAGAGCT 360
DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTAGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGACAGTTTG 420
DB 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 480

DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCCCTTGGATGAGATAGCAGATAGAAA 540
DB 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCCCTTCTGCTGCTAGAGAGGATTTTCTTCTCAGGA 600
DB 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGluValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCATCTTTTCC 660
DB 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTCATCAGCAGATGAAGACTTTCACGTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAGCCTTTCAGAGAAAGGCTCAGGAGAGATCATTTGTTGATGCTGCTCAA 780
DB 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGAGCAGAGGATTTTAAACAGAGCCTTGGCAGTGGCCTCATTTGGAATGAATTCATT 840
DB 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyValAsnCysIle 280
QY 841 TTGATGAACACAGTACATTTGAGTTGTAGCTGACAGATTTACTTGTGAACCTTGGATTCTCA 900
DB 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGGCAGAAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAACGAAAA 960
DB 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTCTTTCAGAAACAGATTTTCAGATATCAGCGTTTTCAGGTTTTCAGTTATGCGA 1020
DB 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGTCCTTCCCTTGGATGCGAGATTTT 1053
DB 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 13
PCT-US02-10824-143
Sequence 143, Application PC/TUS0210824
GENERAL INFORMATION:
APPLICANT: Origene Technologies
TITLE OF INVENTION: Prostate Cancer Expression Profiles
FILE REFERENCE: 9U 206 PCT
CURRENT APPLICATION NUMBER: PCT/US02/10824
CURRENT FILING DATE: 2002-04-08
PRIORITY FILING DATE: 2002-04-08
PRIORITY FILING DATE: 2001-04-06
PRIORITY FILING DATE: 2001-04-06
PRIORITY FILING DATE: 2001-04-06
NUMBER OF SEQ ID NOS: 211
SOFTWARE: PatentIn version 3.1
SEQ ID NO 143
LENGTH: 389
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US02-10824-143

Alignment Scores:
Pred. No.: 9.8e-147 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservativity: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 1 Gaps: 0

US-10-698-228-12 (1-1053) x PCT-US02-10824-143 (1-389)

Qy 1051 TTT 1053
Db 389 Phe 389

Search completed: October 30, 2005, 06:48:03
Job time : 181.25 secs


```
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
Qy 121 CGGTTGTCATCTTCCAACTCCAGTACCTGATATTTGGAAATGTATAACAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Qy 181 GCCTTCCTTCCTGGCAGCAGAGAGGTTGACTTATCAAAAGGATCTCCCTCACTGGAACAAG 240
Db 61 AlaSerPheTrpThrAlaGluValAspLysSerLysAspLeuProHisTrpAsnLys 80
Qy 241 CTTAAACAGATGAGAAGTACTTCACTCTCACATCTTACGCTTTTTCGACGACGAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Qy 301 GGAATCTTAATGAAATTTGGGAGCGCTTTAGTCAGAGGTGCAGGTCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnProGluAla 120
Qy 361 CGCTGTTTCTATGGCTTTCAAATCTCATCGAATCTTCACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Qy 421 CTGTATGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Qy 481 ACCATGCCCTATGTTAAGAAATAAGCAGATGGCGCTTTCAGAGGTGCAGGTCAGAGGCT 540
Db 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Qy 541 TCTACTTTTGGGAAAGAGTGGCTTTCAGAGAGGCTTCTGCTAGAGGAGTTCCTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Qy 601 TCTTTTCTGCTGATATTCTGGCTAAAGAGAGGCTTATGCGAGGACTCACCTTTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Qy 661 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Qy 721 TACTTACTTAATAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATTGCATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Qy 841 TTGATGAAACAGTACATGATGTTGTAGTGCAGAGATTACTTGTGGAACTTGGATCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Qy 901 AAGTTTTCAGGAGAGAAATCCTTTTGAATTTATGAAACATTTCTTTAGAGGAAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Qy 961 ACAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGGCTTTTTCAGGTATGCGAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Qy 1021 ACAGATAACGTTCTACCTTGGATGCAATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 2

```
PCT-US05-21650-34
; Sequence 34, Application PC/TUS0521650
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: MIGFS AS MODIFIERS OF THE IGF PATHWAY AND METHODS OF USE
```

```
; FILE REFERENCE: EX05-017
; CURRENT APPLICATION NUMBER: PCT/US05/21650
; CURRENT FILING DATE: 2005-06-27
; PRIOR APPLICATION NUMBER: 60/581,689
; PRIOR FILING DATE: 2004-06-21
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; PCT-US05-21650-34

Alignment Scores:
Pred. No.: 1,19e-180 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 1 Gaps: 0

US-10-698-228-12 (1-1053) x PCT-US05-21650-34 (1-351)
Qy 1 ATGGGCGACCCGGAAGCGCGGCGCTGGATCAGATGAGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Qy 61 GACACCAACCAAGTGAATTAAGTCAANTGAAGAGCCACTCCTAAGAAAGTTCGCG 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
Qy 121 CGGTTGCTCATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATAACAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Qy 181 GCCTTCTTCTGAGCAGCAGAGAGGTTGACTTATCAAAAGGATCTCCCTCAGTGAACAAG 240
Db 61 AlaSerPheTrpThrAlaGluValAspLysSerLysAspLeuProHisTrpAsnLys 80
Qy 241 CTTAAACAGATGAGAAGTACTTCACTCTCACATCTTACGCTTTTTCGACGACGAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Qy 301 GGAATCTTAATGAAATTTGGGAGCGCTTTAGTCAGAGGTGCAGGTCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnProGluAla 120
Qy 361 CGCTGTTTCTATGGCTTTCAAATCTCATCGAATCTTCACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Qy 421 CTGTATGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Qy 481 ACCATGCCCTATGTTAAGAAATAAGCAGATGGCGCTTTCAGAGGTGCAGGTCAGAGGCT 540
Db 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Qy 541 TCTACTTTTGGGAAAGAGTGGCTTTCAGAGAGGCTTCTGCTAGAGGAGTTCCTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Qy 601 TCTTTTCTGCTGATATTCTGGCTAAAGAGAGGCTTATGCGAGGACTCACCTTTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Qy 661 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Qy 721 TACTTACTTAATAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATTGCATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Qy 841 TTGATGAAACAGTACATGATGTTGTAGTGCAGAGATTACTTGTGGAACTTGGATCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Qy 901 AAGTTTTCAGGAGAGAAATCCTTTTGAATTTATGAAACATTTCTTTAGAGGAAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Qy 961 ACAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGGCTTTTTCAGGTATGCGAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Qy 1021 ACAGATAACGTTCTACCTTGGATGCAATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

```
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluLeuValAlaValLys 260
QY 781 ATTGACGAGGTTTTTAACAGAGCCTTGCAGAGTGGCTCATTTGGAATGATTCATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGCAGAAAATCCTTTGATTTATGCGAAACATCTTTTAGAGGAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTGAGAAACAGATTTCAGAGTATCAGCGTTTTTCAGTTATGGCAGAAACC 1020
Db 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGTTCTTCACTTGGATGCGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAlaAspPhe 351

RESULT 3
US-10-990-328A-7547
; Sequence 7547, Application US/10990328A
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328A
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7547
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328A-7547

Alignment Scores:
Pred. No.: 1.19e-180 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 6 Gaps: 0

US-10-698-228-12 (1-1053) x US-10-990-328A-7547 (1-351)
QY 1 ATGGGCGACCCGGAAGCCGAGCGCGCTGGATCAGATGAGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAGTGAATAAGTCAATGAGAGCCACTCTTAAGAAGAGTTCTCGC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
QY 121 CGTTTGTTCATCTTCCCAATCCAGTACCTCGATATTGGAAAATGATATAACAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTCTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAAACAAG 240
Db 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAGCAGATCAGAGTACTTCTCTCATCTTAGCCCTTTTTCAGCCAGTGTAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
Db 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
```

```
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCAGAAATGTTCTACCTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTATTAAATGCAATTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATGGCCCTTCCGATGCATGATGAGCAGATGAAA 540
Db 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAAGAGTGGCTGCTGCTGCTGTAGAGAGGATTTTCTTCTCAGAGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTCTGCTATATTCTGGCTAAAGAAGAGAGGTCTTATGCCAGGAGCTCACCTTTTCC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAATCATCAGCAGAGATGAAGACTTCACCTGTGTGACTTTGCTGCTGATGTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATAAGCCTTCAGAGAAAGGTCAGGAGAGATCATTTGATGCTGCTCTCAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGAGCAGGAGTTTTTAACAAGACCTTGCAGCTTGGCCCTCATTTGGAATGAATGCAATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACAGTACATTGAGTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATTCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGTTTTTTCAGCAGAAAATCCTTTGATTTATGGAACCATTTCTTTAGAAGGAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTGAGAAACAGATTTCAGAGTATCAGCGTTTTTGCAGTTTATGCGAGAAACC 1020
Db 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTTCTTCACTTGGATGCGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAlaAspPhe 351

RESULT 4
US-11-088-686-15
; Sequence 15, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-15

Alignment Scores:
Pred. No.: 1.19e-180 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
```

```

Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 98.65%                Indels: 0
DB: 7                               Gaps: 0

US-10-698-228-12 (1-1053) x US-11-088-686-15 (1-351)

QY 1 ATGGGCGACCCGGAAGGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACACGAAGTGAATAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
DB 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGGTTTGTTCATCTTCCATCCAGTACCTGATATTTGGAATGTATTAACAGGCACAG 180
DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTCTGACAGCAGAGAGGTGTGACTTATCAAGGATCTCCTCACTGGAACAAG 240
DB 61 AlaSerPheTyrThrAlaGluValAspLeuSerLysAspLeuProHisTyrPheLys 80
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCAGATCTTACGCTTTTTCAGCAGTAT 300
DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAlaSerAsp 100
QY 301 GGAATTTGAATGAATAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnProGluAla 120
QY 361 CGCTGTTTCTATGCTTCAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
DB 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGATCCAAAGAAAGGGAATTTTATTAATGCAATGAA 480
DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACATGCGCTATGTTAAGAAAAGCAGATTTGGCCCTTGGATGGATGAGATAGAAAA 540
DB 161 ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrPheAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTGTCTGTAGAGAGTTTCTTCTCAGGA 600
DB 181 SerThrPheGlyGluArgValAlaIlePheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTGCTATATTCGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
DB 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTCATCAGCAGATGAGGACTTCACTGTGACTTTGCTTGGCTGTGATGTTCAA 720
DB 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTGATGCTGTCAAA 780
DB 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGAGCAGGAGTTTTTAAACAGAGCCTTGCAGAGTTGGCTCATTTGGAATGAATTCATT 840
DB 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACACGATGATGATTTGATGCTGACAGATTAATCTGTGGAACCTTGGATTTCA 900
DB 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTTCGCGACAAAATCCTTTTGAATTTTATGAAAACATTTCTTTAGAGAAAA 960
DB 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTTCAGAAAACGAGTTTCAGAGTATCAGCGTTTTCAGATTATGGCAGAAAC 1020
DB 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340

US-10-698-228-12 (1-1053) x PCT-US05-09639-17 (1-351)

QY 1 ATGGCGGACCCGGAAGGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACACGAAGTGAATAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
DB 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGGTTTGTTCATCTTCCATCCAGTACCTGATATTTGGAATGTATTAACAGGCACAG 180
DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTCTGACAGCAGAGAGGTGTGACTTATCAAGGATCTCCTCACTGGAACAAG 240
DB 61 AlaSerPheTyrThrAlaGluValAspLeuSerLysAspLeuProHisTyrPheLys 80
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCAGATCTTACGCTTTTTCAGCAGTAT 300
DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAlaSerAsp 100
QY 301 GGAATTTGAATGAATAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnProGluAla 120
QY 361 CGCTGTTTCTATGCTTCAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
DB 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGATCCAAAGAAAGGGAATTTTATTAATGCAATGAA 480
DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACATGCGCTATGTTAAGAAAAGCAGATTTGGCCCTTGGATGGATGAGATAGAAAA 540
DB 161 ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrPheAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTGTCTGTAGAGAGTTTCTTCTCAGGA 600
DB 181 SerThrPheGlyGluArgValAlaIlePheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTGCTATATTCGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
DB 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTCATCAGCAGATGAGGACTTCACTGTGACTTTGCTTGGCTGTGATGTTCAA 720
DB 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTGATGCTGTCAAA 780
DB 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGAGCAGGAGTTTTTAAACAGAGCCTTGCAGAGTTGGCTCATTTGGAATGAATTCATT 840
DB 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACACGATGATGATTTGATGCTGACAGATTAATCTGTGGAACCTTGGATTTCA 900
DB 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTTCGCGACAAAATCCTTTTGAATTTTATGAAAACATTTCTTTAGAGAAAA 960
DB 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTTCAGAAAACGAGTTTCAGAGTATCAGCGTTTTCAGATTATGGCAGAAAC 1020
DB 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
```

QY 601 TCTTTGCTGCTATATTCTGGCTAAAGAGAGAGGCTTATGCGGAGACTCACTTTTCC 660
DB 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACATCATGACGAGATGAGAGCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
DB 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAAGCTTCAAGAGAAAGGCTCAGGAGATCATTTGATGCTGTCMAA 780
DB 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGACGAGAGTATTTTAAACAGAGCTTGCACAGTGGCTCATTTGGAATGGAATTGCA 840
DB 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACAGTACATGAGTTGTAGTGCAGAGATTAATCTGTGAACTTGGATTTCTCA 900
DB 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGTTTTCAGGACAGAAATCTTTGATTGTTTATGGAACATTTCTTTAGAGAA 960
DB 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTCTTTCAGAGAAAGGTTTCAGAGTATCAGCGTTTTCAGCTTATGGCAGAAAC 1020
DB 321 ThrAsnPheGlnLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053
DB 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 6

US-11-088-686-17

; Sequence 17, Application US/11088686

; GENERAL INFORMATION:

; APPLICANT: Yen, Yun

; TITLE OF INVENTION: DRUG SCREENING

; FILE REFERENCE: 14037-004001

; CURRENT APPLICATION NUMBER: US/11/088,686

; CURRENT FILING DATE: 2005-03-23

; PRIOR APPLICATION NUMBER: US 60/556,836

; PRIOR FILING DATE: 2004-03-25

; NUMBER OF SEQ ID NOS: 59

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 17

; LENGTH: 351

; TYPE: PRT

; ORGANISM: Homo sapiens

US-11-088-686-17

Alignment Scores:

Pred. No.: 3,11e-180 Length: 351

Score: 1817.00 Matches: 350

Percent Similarity: 100.00% Conservative: 1

Best Local Similarity: 99.72% Mismatches: 0

Query Match: 98.43% Indels: 0

DB: 7 Gaps: 0

US-10-698-228-12 (1-1053) x US-11-088-686-17 (1-351)

QY 1 ATGGGCGACCCGGAAGGCGGAGCGCGGCTGGATCAGATGAGAGATCATTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAAGCAAGTAAATCAATGAAGCCACTCTTAAGAAGAGTCTCCG 120
DB 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGGTTTGTCTATCTTCCAAATCAGTACCTGATATTTTGGAAATGTATATAACAGGCACAG 180
DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60


```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-19

Alignment Scores:
Pred. No.:      3,95e-180      Length:      351
Score:          1816.00      Matches:      350
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.72%      Mismatches: 0
Query Match:      98.37%      Indels:      0
DB:              1          Gaps:      0

US-10-698-228-12 (1-1053) x PCT-US05-09639-19 (1-351)

QY 1 ATGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACCAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
QY 121 CGTTTGTCTATCTTCCATCCAGTACCTGCTATATTGGAAAATGTATAAACAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTCTGACAGCAGAGAGGTGTGACTTATCAAAAGGATCTCCCTCACTGGAACAAG 240
Db 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCAGATCTTACGCTTTTTCGACCGCAGTGAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 CGAATTGTAAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTrpGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTATGAATGTAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAGAAAGCAGATTGGCCCTTCGATGGATACGAGTAAAA 540
Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGGAGTCTTCTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTGCTATATTCTGGCTAAAGAGAGAGGTCCTTATGTCAGGACTCACTTTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAAGCTTTCAGAAAGAAAGGTCAGGAGATCATCTGTGATGCTGTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGAGCAGGAGTTTTAAACAGAAAGCTTGCAGAGTGTGCTCATTTGGAATGCAATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
```

```
QY 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGTTTTTTCAGCAGCAAAATCCTTTTGTATTTTATGAAAACATTCTTTTGAAGAGAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAATTTCTTTCAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGTGTATCGCAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAAACGCTTCTTCCACTTTGGATGACAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 8
US-11-088-686-19
; Sequence 19, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-19

Alignment Scores:
Pred. No.:      3,95e-180      Length:      351
Score:          1816.00      Matches:      350
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.72%      Mismatches: 0
Query Match:      98.37%      Indels:      0
DB:              7          Gaps:      0

US-10-698-228-12 (1-1053) x US-11-088-686-19 (1-351)

QY 1 ATGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACCAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
QY 121 CGTTTGTCTATCTTCCATCCAGTACCTGCTATATTGGAAAATGTATAAACAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTCTGACAGCAGAGAGGTGTGACTTATCAAAAGGATCTCCCTCACTGGAACAAG 240
Db 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCAGATCTTACGCTTTTTCGACCGCAGTGAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 CGAATTGTAAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTrpGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTATGAATGTAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAGAAAGCAGATTGGCCCTTCGATGGATACGAGTAAAA 540
Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGGAGTCTTCTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTGCTATATTCTGGCTAAAGAGAGAGGTCCTTATGTCAGGACTCACTTTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAAGCTTTCAGAAAGAAAGGTCAGGAGATCATCTGTGATGCTGTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGAGCAGGAGTTTTAAACAGAAAGCTTGCAGAGTGTGCTCATTTGGAATGCAATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
```

```
QY 421 CTGATAGACATTACATCAGAGATCCCAAGAAAGGGAATTTTATTATTAATGCAATTGAA 480
Db 141 LeuIleAspThrTyrIleArgaspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAGACAGATTGGCCCTCGGATGGATACAGATAGAAA 540
Db 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGCTTGTCTGTCTAGAGAGAGTCTTCTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTCTGGCTAAAGACAGAGTCTTATGCCAGGACTCTCTTTTCC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTCATCAGACAGATGAGGACTTCTACTGTGACTTGTCTGCTGCTGATTTCCAA 720
Db 221 AsnGluLeuIleSerArgaspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAAGCTTTCAGAAAGGCTTCCAGGAGATCATTTCTGATGCTGTCAA 780
Db 241 TyrLeuValAsnLysProSerGluArgValArgGluIleValAspAlaValLys 260
QY 781 ATTGACGAGGAGTCTTTTAAAGAGCTTCCAGGCTTGGCTCATTTGGAATGAATTCATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACAGTACATTGAGTTGTAGCTGACAGATTAATTGTGGAATTGGATTTCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGGCAGAAATCTTTTGAATTTATGGAACATTTCTTAGAAGGAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATTGACAGAAACC 1020
Db 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

```
RESULT 9
PCT-US05-09639-21
; Sequence 21, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen. Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004WO1
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-21
```

```
Alignment Scores:
Pred. No.: 8,12e-180 Length: 351
Score: 1813.00 Matches: 349
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.43% Mismatches: 0
Query Match: 98.21% Indels: 0
DB: 1 Gaps: 0
```

US-10-698-228-12 (1-1053) x PCT-US05-09639-21 (1-351)

RESULT 10

```
QY 1 ATGGCGGACCCGGAAGGCCGAGCCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaIleGlyLeuAspGlnAspGluArgSerSerSer 20
QY 61 GACACCAACCAAGTGAATAAGTCAATCAAGAGCCACTCCTTAAGAAAGATTCTCCG 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
QY 121 CGGTTTGTCTATCTTCCAAATCCAGTACCTCATATTGGAAAAATGTATAAACAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTTGGCAGCAGAGAGGTTGACTTATCAAGAGTCTCCCTCCTCAGTGGAAACAG 240
Db 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAGCAGATCAGAAAGTACTTCTCATCTCTCATCTTAGCCTTTTGTGACGCCAGTGAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTTCTAAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATCGCTTCAAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPhePheGlyPheGlnIleLeuIleGluAsnValHisSerGluMetPheSerLeu 140
QY 421 CTGATAGACATTACATCAGAGATCCCAAGAAAGGGAATTTTATTATTAATGCAATTGAA 480
Db 141 LeuIleAspThrTyrIleArgaspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAGACAGATTGGCCCTTGGATGGATGAGATAGAAA 540
Db 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGCTTGTCTCTGTAGAGGAGTTTCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTCTGGCTAAAGAGAGGCTTCTATGCCAGGACTCACATTTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCTACTGTGACTTGTGCTGCTGATTTCCAA 720
Db 221 AsnGluLeuIleSerArgaspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAAGCTTTCAGAAAGGCTTCCAGGAGATCATTTGTGATGCTGTCAA 780
Db 241 TyrLeuValAsnLysProSerGluArgValArgGluIleValAspAlaValLys 260
QY 781 ATTGACGAGGAGTCTTTTAAAGAGGCTTCCAGGCTTGGCTCATTTGGAATGAATTCATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACAGTACATTGAGTTGTAGCTGACAGATTAATTGTGGAATCTGGATTTCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGGCAGAAATCTTTTGAATTTATGGAACATTTCTTAGAAGGAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGGCTTTCAGTATTGACAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

```
US-11-088-686-21
; Sequence 21, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-21

Alignment Scores:
Pred. No.:      8,12e-180      Length:      351
Score:          1813.00      Matches:      349
Percent Similarity: 100.00%      Conservative: 2
Best Local Similarity: 99.43%      Mismatches: 0
Query Match:      98.21%      Indels:      0
DB:              7          Gaps:      0

US-10-698-228-12 (1-1053) x US-11-088-686-21 (1-351)
QY 1 ATGGCGGACCCGGAAGCGCGGCGGCTGGATCAGAGTCAGAGATCATCTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAAGTGAATGAAGTCAATGAAGCCACTCTAAGAAGAGTTCTCGC 120
DB 21 AspThrAsnGluSerGluIleYsSerAsnGluProLeuLeuArgLysSerArg 40
QY 121 CGTTTGTTCATCTTCCCAATCCAGTACCTCTATATTTGGAATGTATATAACAGGCACAG 180
DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTTCTGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGACAACAG 240
DB 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAGCAGATCAGAGTACTTCTCATCTCTCATCTTAGCCTTTTTCAGCCAGTAT 300
DB 81 LeuLysAlaAspGluYsTrpPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATGTGAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCAGAGGCT 360
DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
DB 121 ArgCysPhePheGlyPheGlnIleLeuIleGluAsnValHisSerGluMetPheSerLeu 140
QY 421 CTGATAGACACTTACATCAGATCCCAAGAAAGGAAATTTTATTTAATGCAATGAA 480
DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTCGGCCTTCGATGGATAGCAGATAGAAA 540
DB 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGGAAGAGTGTGGCTTTGTGCTGTAGAGAGGAGTTTCTTCTCAGGA 600
DB 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTGCTATATCTGGCTAAAGAGAGAGGTCTTATCCAGGACTCACTTTTCC 660
DB 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACATCAGCAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
```

```
DB 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGCTCAAA 780
DB 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGACGAGGAGTGTAAACAGAAAGCCTTCAGAGTGGCCTCATTTGGAATGAATTCGATT 840
DB 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
DB 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTCAGCGCAGAAAATCCTTTTGTATTTTATGAAAACATTTCTTTAGAGGAAAA 960
DB 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGTATGGCAGAAACC 1020
DB 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTCTTCACCTTGGATGAGATT 1053
DB 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 11
PCT-US05-09639-23
; Sequence 23, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-23

Alignment Scores:
Pred. No.:      1.03e-179      Length:      351
Score:          1812.00      Matches:      349
Percent Similarity: 100.00%      Conservative: 2
Best Local Similarity: 99.43%      Mismatches: 0
Query Match:      98.16%      Indels:      0
DB:              1          Gaps:      0

US-10-698-228-12 (1-1053) x PCT-US05-09639-23 (1-351)
QY 1 ATGGCGGACCCGGAAGCGCGGCGGCTGGATCAGAGTCAGAGATCATCTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAAGTGAATGAAGTCAATGAAGCCACTCTCAAGAGAGTTCCTCGC 120
DB 21 AspThrAsnGluSerGluIleYsSerAsnGluProLeuLeuArgLysSerArg 40
QY 121 CGTTTGTTCATCTTCCCAATCCAGTACCTCTATATTTGGAATGTATATAACAGGCACAG 180
DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTTCTGACAGCAGAGGTGACTTATCAAGGATCTCCCTCACTGGACAACAG 240
DB 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAGCAGATCAGAGTACTTCTCATCTCTCATCTTAGCCTTTTTCAGCCAGTAT 300
```

```
Db      81  LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Qy      301  CGAAATTGTAATAAGAAATTTGGTGGAGCCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
Db      101  GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Qy      361  CGCTGTTCTATGCGCTTTCAAAATTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 420
Db      121  ArgCysPhePheGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTrpSerLeu 140
Qy      421  CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db      141  LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Qy      481  ACCATGCCCTATGTTAAGAAAGACAGATTGGCGCTTCGGATGGATACAGATAGAAA 540
Db      161  ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Qy      541  TCTACTTTTGGGAAAGAGTGGCTTTGCTGTGTAGAGGAGTTCCTTCTCAGGA 600
Db      181  SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
Qy      601  TCTTTTGTCTATATTCGGCTTAAGAGAGAGGTCTTATGCGAGACTCACTTTTTC 660
Db      201  SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Qy      661  ANTGAACATCATCAGAGATCAGAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db      221  AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Qy      721  TACTTAGTAAATAAGCCTTCAGAAAGAGGTTCAGGAGATCATCTTCATGCTGTCAAA 780
Db      241  TyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAlaValLys 260
Qy      781  ATTGACGAGGAGTTTAAACAGAGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCATT 840
Db      261  IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Qy      841  TTGATGAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGGAACTTGGATTCTCA 900
Db      281  LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Qy      901  AAGGTTTTTCAGGCAGAAATCCTTTGATTTATGGAACATTTCTTAGAAGAAA 960
Db      301  LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Qy      961  ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCGAGTTATGGCAGAAAC 1020
Db      321  ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Qy      1021  ACAGATAACGCTTTCACCTTGGATCGAGATTTT 1053
Db      341  ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 12

```
PCT-US05-09639-25
; Sequence 25, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004WO1
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-25
```

```
Alignment Scores:
Pred. No.:      1.03e-179      Length:      351
Score:          1812.00      Matches:      349
Percent Similarity: 100.00%      Conservative: 2
Best Local Similarity: 99.43%      Mismatches: 0
Query Match:      98.16%      Indels: 0
DB:              1          Gaps: 0
```

US-10-698-228-12 (1-1053) x PCT-US05-09639-25 (1-351)

```
Qy      1  ATGGGCGACCGCGAAAGCGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db      1  MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Qy      61  GACACCAACCAAGTGAATTAAGTCAATCAAGAGCCACTCCTTAAGAAAGTTCTCG 120
Db      21  AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
Qy      121  CGGTTTGTCTATCTTCCAAATCCAGTACCTCATATTTGAAAAATGTATAACAGGCACAG 180
Db      41  ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Qy      181  GCTTCTCTCTGACAGCAGAGAGGTTCACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db      61  AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Qy      241  CTTAAGCAGATGAAAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCGACCGCATGTAT 300
Db      81  LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Qy      301  GGAATTTCTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
Db      101  GlyIleValAsnGluAsnLeuValArgPheSerGlnGluValGlnValProGluAla 120
Qy      361  CGCTGTTTCTATGCTTTCAAATTTCTCATCAGAGATGTTCACTCAGAGATGTACAGTTTG 420
Db      121  ArgCysPheTrpGlyPheGlnIleLeuIleGluAsnValHisSerGluMetPheSerLeu 140
Qy      421  CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db      141  LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Qy      481  ACCATGCCCTATGTTAAGAAAGACAGATTGGCGCTTCGGATGGATACAGATAGAAA 540
Db      161  ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Qy      541  TCTACTTTTGGGAAAGAGTGGCTTTGCTGTGTAGAGGAGTTCCTTCTCTCAGGA 600
Db      181  SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
Qy      601  TCTTTTGTCTATATTCGGCTTAAGAGAGAGGTCTTATGCCAGGAGCTCACTTTTTC 660
Db      201  SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Qy      661  ANTGAACATCATCAGAGATGAAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db      221  AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Qy      721  TACTTAGTAAATAAGCCTTCAGAAAGAGGTTCAGGAGATCATTTGTTGATGCTCTCAAA 780
Db      241  TyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAlaValLys 260
Qy      781  ATTGACGAGGAGTTTAAACAGAGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCATT 840
Db      261  IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Qy      841  TTGATGAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGGAACTTGGATTCTCA 900
Db      281  LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Qy      901  AAGGTTTTTCAGGCAGAAATCCTTTGATTTATGGAACATTTCTTAGAAGAAA 960
```

Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTCTTTGAGAACGAGTTTCAGAGTATCAGCGTTTTCAGTGTATGCGCAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTCTTCCACCTTGGATCGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 13
US-11-088-686-23
; Sequence 23, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-23

Alignment Scores:
Pred. No.: 1.03e-179 Length: 351
Score: 1812.00 Matches: 349
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.43% Mismatches: 0
Query Match: 98.16% Indels: 0
DB: 7 Gaps: 0

US-10-698-228-12 (1-1053) x US-11-088-686-23 (1-351)

QY 1 ATGGCGACCCGGAAGAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAAGAAAGTAAGTCAATGAAGACCACTCTAAGAAAGAGTTCTCCG 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGTTTGTTCATCTTCCCAATCCAGTACCCTGATATTTGAAAATGTATAAAGCAGACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTTCTGACACGACGAGAGGTGTGACTATCAAGGATCTCCCTCACTGGAAACAAG 240
Db 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProIleThrAsnLys 80
QY 241 CTAAAGCAGATGAGAGTACTTCTCTCATCTTACCTTACCTTTTTCAGCCAGTGAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTCCAAATTCATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPhePheGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTrpSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCGCTTCGGATGGATAGCAGATAGAAA 540

Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAAGAGTGGCTTTGCTGCTCTAGAGGAGTTTCTTCTCAGCA 600
Db 181 SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACATTTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACTCATCAGCAGAGATGAAGACTTCATCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 221 AsnGluIleSerArgGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAAGCTTCAGAGAAAGGCTCAGGAGAGATCATTTGTTGCTGCTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGACGAGGAGTTTAAACAGAAAGCCTTGCAGATTGGCCCTCATTTGGAATGAATTCAT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAAACAGTACATTGAGTTGTAGCTGACAGATTAATTGTGGAACCTTGGATTCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGGCAGAAATCCTTTTGTATTTATGGAACATTTCTTTAGAGGAAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTCTTTCAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGGTTTATGCGCAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTCTTCCACCTTGGATCGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 14

US-11-088-686-25
; Sequence 25, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-25

Alignment Scores:
Pred. No.: 1.03e-179 Length: 351
Score: 1812.00 Matches: 349
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.43% Mismatches: 0
Query Match: 98.16% Indels: 0
DB: 7 Gaps: 0

US-10-698-228-12 (1-1053) x US-11-088-686-25 (1-351)

QY 1 ATGGCGACCCGGAAGAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAAGAAAGTAAATGTAAGTCAAAATGAAGAGCCACTCTTCAAGAAAGAGTTCTCGC 120

```
Db 21 AspThrAsnGluSerGluLeuIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
QY 121 CGGTTTGTCACTTTCCCAATCCAGTACCTGATATTTGGAAATGATATAACAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTTGACACAGACAGAGGTTGACTTATCAAGGATCTCCCTCAGTGGAAACAG 240
Db 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrPheAsnLys 80
QY 241 CTTAAGCAGATGAGAACTTCTCATCTCTCATCTTAGCCCTTTTTCAGCCAGCTGAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTTGTAATGAAATTTGGTGAGCGCTTGTAGTCAGAGGTGCAGCTTCCAGAGGCT 360
Db 101 GlyIleValAsnGlnAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAAATGTTCTCACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyrPheGlnIleLeuIleGluAsnValHisSerGluMetPheSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTATGCAATTTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTCCGATGATACGATAGAAA 540
Db 161 ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrPheAlaAspArgLys 180
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTAGAGAGGTTTCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaValAlaGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATCTGCTGCTAAAGAGAGGCTTTATGCCAGGACTCACTTTTTC 660
Db 201 SerPheAlaAlaIlePheTyrLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACTCATCAGAGATGAGGACTTCACTGTGACTTGTGCTGCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATTAAGCTTTCAGAAAGAGGCTCAGGAGATCATTTGTGATGCTGCTCAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGACGAGGAGTTTAAACAGAGCCCTTGCAGTGGCCCTCATTTGGAATGCAATTCAT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAAACAGTACATGAGTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATTTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGGCAGAGAAATCTTTTGTATTTTATGGAACATTTCTTTAGAGGAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTCTTTGAGAACAGGTTTTCAGAGTATCAGCGTTTTCAGAGTATGCGAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATACGCTTTCACCTTGGATGCGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 15

```
PCT-US05-09639-27
; Sequence 27, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004WO1
```

```
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-27
```

Alignment Scores:

```
Pred. No.: 1,31e-179 Length: 351
Score: 1811.00 Matches: 349
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.43% Mismatches: 0
Query Match: 98.10% Indels: 0
DB: 1 Gaps: 0
```

US-10-698-228-12 (1-1053) x PCT-US05-09639-27 (1-351)

```
QY 1 ATGGGCGACCCGGAAGCGCGGCGCTGGATCAGGATGAGAGATCATCTTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAAGTCAATTAAGTCAATGAAGAGCCACTCCTAAGAAAGATTCGCG 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGGTTTGTCTATCTTCCAAATCCAGTACCTGATATTGGAAATGTATAACAGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTGTCAGCAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCAGTGGAAACAAG 240
Db 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrPheAsnLys 80
QY 241 CTTAAGCAGATGAGAACTTCTCATCTCTCATCTTAGCCCTTTTTCAGCCAGCTGAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTTGTAATGAAATTTGGTGAGCGCTTTGTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db 101 GlyIleValAsnGlnAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAAATGTTCTCACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyrPheGlnIleLeuIleGluAsnValHisSerGluMetTyrPheSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTATGCAATTTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTCCGATGATACGATAGAAA 540
Db 161 ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrPheAlaAspArgLys 180
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAAGAGGTTTCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaValAlaGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATCTGCTGCTAAAGAGAGGCTTTATGCCAGGACTCACTTTTTC 660
Db 201 SerPheAlaAlaIlePheTyrLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACTCATCAGAGATGAGGACTTCACTGTGACTTGTGCTGCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATTAAGCTTTCAGAAAGAGGCTCAGGAGATCATTTGTGATGCTGCTCAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2005, 23:48:01 ; Search time 3214.5 Seconds
(without alignments)
13372.738 Million cell updates/sec

Title: US-10-698-228-12

Perfect score: 1053

Sequence: 1 atggggaccgcgaaagcc.....tcaccttgatgagatttt 1053

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 45554873 seqs, 20411521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents NA Main:*

- 1: /cgn2_6/ptodata/1/pna/PCTUS1 COMB.seq:*
- 2: /cgn2_6/ptodata/1/pna/PCTUS2 COMB.seq:*
- 3: /cgn2_6/ptodata/1/pna/PCTUS COMB.seq:*
- 4: /cgn2_6/ptodata/1/pna/US06 COMB.seq:*
- 5: /cgn2_6/ptodata/1/pna/US07 COMB.seq:*
- 6: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*
- 7: /cgn2_6/ptodata/1/pna/US081 COMB.seq:*
- 8: /cgn2_6/ptodata/1/pna/US082 COMB.seq:*
- 9: /cgn2_6/ptodata/1/pna/US083 COMB.seq:*
- 10: /cgn2_6/ptodata/1/pna/US084 COMB.seq:*
- 11: /cgn2_6/ptodata/1/pna/US085 COMB.seq:*
- 12: /cgn2_6/ptodata/1/pna/US086 COMB.seq:*
- 13: /cgn2_6/ptodata/1/pna/US087 COMB.seq:*
- 14: /cgn2_6/ptodata/1/pna/US088 COMB.seq:*
- 15: /cgn2_6/ptodata/1/pna/US089 COMB.seq:*
- 16: /cgn2_6/ptodata/1/pna/US090 COMB.seq:*
- 17: /cgn2_6/ptodata/1/pna/US091 COMB.seq:*
- 18: /cgn2_6/ptodata/1/pna/US092A COMB.seq:*
- 19: /cgn2_6/ptodata/1/pna/US092B COMB.seq:*
- 20: /cgn2_6/ptodata/1/pna/US093A COMB.seq:*
- 21: /cgn2_6/ptodata/1/pna/US093B COMB.seq:*
- 22: /cgn2_6/ptodata/1/pna/US094 COMB.seq:*
- 23: /cgn2_6/ptodata/1/pna/US095A COMB.seq:*
- 24: /cgn2_6/ptodata/1/pna/US095B COMB.seq:*
- 25: /cgn2_6/ptodata/1/pna/US095C COMB.seq:*
- 26: /cgn2_6/ptodata/1/pna/US095D COMB.seq:*
- 27: /cgn2_6/ptodata/1/pna/US096A COMB.seq:*
- 28: /cgn2_6/ptodata/1/pna/US096B COMB.seq:*
- 29: /cgn2_6/ptodata/1/pna/US096C COMB.seq:*
- 30: /cgn2_6/ptodata/1/pna/US096D COMB.seq:*
- 31: /cgn2_6/ptodata/1/pna/US096E COMB.seq:*
- 32: /cgn2_6/ptodata/1/pna/US097A COMB.seq:*
- 33: /cgn2_6/ptodata/1/pna/US097B COMB.seq:*
- 34: /cgn2_6/ptodata/1/pna/US097C COMB.seq:*
- 35: /cgn2_6/ptodata/1/pna/US098A COMB.seq:*
- 36: /cgn2_6/ptodata/1/pna/US098B COMB.seq:*
- 37: /cgn2_6/ptodata/1/pna/US098C COMB.seq:*
- 38: /cgn2_6/ptodata/1/pna/US098D COMB.seq:*
- 39: /cgn2_6/ptodata/1/pna/US099A COMB.seq:*
- 40: /cgn2_6/ptodata/1/pna/US099B COMB.seq:*
- 41: /cgn2_6/ptodata/1/pna/US099C COMB.seq:*
- 42: /cgn2_6/ptodata/1/pna/US099D COMB.seq:*
- 43: /cgn2_6/ptodata/1/pna/US099E COMB.seq:*

- 44: /cgn2_6/ptodata/1/pna/US099F COMB.seq:*
- 45: /cgn2_6/ptodata/1/pna/US099G COMB.seq:*
- 46: /cgn2_6/ptodata/1/pna/US100A COMB.seq:*
- 47: /cgn2_6/ptodata/1/pna/US100B COMB.seq:*
- 48: /cgn2_6/ptodata/1/pna/US101A COMB.seq:*
- 49: /cgn2_6/ptodata/1/pna/US101B COMB.seq:*
- 50: /cgn2_6/ptodata/1/pna/US102A COMB.seq:*
- 51: /cgn2_6/ptodata/1/pna/US102B COMB.seq:*
- 52: /cgn2_6/ptodata/1/pna/US103A COMB.seq:*
- 53: /cgn2_6/ptodata/1/pna/US103B COMB.seq:*
- 54: /cgn2_6/ptodata/1/pna/US104A COMB.seq:*
- 55: /cgn2_6/ptodata/1/pna/US104B COMB.seq:*
- 56: /cgn2_6/ptodata/1/pna/US105A COMB.seq:*
- 57: /cgn2_6/ptodata/1/pna/US105B COMB.seq:*
- 58: /cgn2_6/ptodata/1/pna/US106A COMB.seq:*
- 59: /cgn2_6/ptodata/1/pna/US107A COMB.seq:*
- 60: /cgn2_6/ptodata/1/pna/US107B COMB.seq:*
- 61: /cgn2_6/ptodata/1/pna/US107C COMB.seq:*
- 62: /cgn2_6/ptodata/1/pna/US107D COMB.seq:*
- 63: /cgn2_6/ptodata/1/pna/US108A COMB.seq:*
- 64: /cgn2_6/ptodata/1/pna/US108B COMB.seq:*
- 65: /cgn2_6/ptodata/1/pna/US109A COMB.seq:*
- 66: /cgn2_6/ptodata/1/pna/US109B COMB.seq:*
- 67: /cgn2_6/ptodata/1/pna/US109C COMB.seq:*
- 68: /cgn2_6/ptodata/1/pna/US110 COMB.seq:*
- 69: /cgn2_6/ptodata/1/pna/US6000 COMB.seq:*
- 70: /cgn2_6/ptodata/1/pna/US6001 COMB.seq:*
- 71: /cgn2_6/ptodata/1/pna/US6002 COMB.seq:*
- 72: /cgn2_6/ptodata/1/pna/US6003 COMB.seq:*
- 73: /cgn2_6/ptodata/1/pna/US6004 COMB.seq:*
- 74: /cgn2_6/ptodata/1/pna/US6005 COMB.seq:*
- 75: /cgn2_6/ptodata/1/pna/US6006 COMB.seq:*
- 76: /cgn2_6/ptodata/1/pna/US6007 COMB.seq:*
- 77: /cgn2_6/ptodata/1/pna/US6008 COMB.seq:*
- 78: /cgn2_6/ptodata/1/pna/US6009 COMB.seq:*
- 79: /cgn2_6/ptodata/1/pna/US6010 COMB.seq:*
- 80: /cgn2_6/ptodata/1/pna/US6011 COMB.seq:*
- 81: /cgn2_6/ptodata/1/pna/US6012 COMB.seq:*
- 82: /cgn2_6/ptodata/1/pna/US6013 COMB.seq:*
- 83: /cgn2_6/ptodata/1/pna/US6014 COMB.seq:*
- 84: /cgn2_6/ptodata/1/pna/US6015 COMB.seq:*
- 85: /cgn2_6/ptodata/1/pna/US6016 COMB.seq:*
- 86: /cgn2_6/ptodata/1/pna/US6017 COMB.seq:*
- 87: /cgn2_6/ptodata/1/pna/US6018 COMB.seq:*
- 88: /cgn2_6/ptodata/1/pna/US6019 COMB.seq:*
- 89: /cgn2_6/ptodata/1/pna/US6020 COMB.seq:*
- 90: /cgn2_6/ptodata/1/pna/US6021 COMB.seq:*
- 91: /cgn2_6/ptodata/1/pna/US6022 COMB.seq:*
- 92: /cgn2_6/ptodata/1/pna/US6023A COMB.seq:*
- 93: /cgn2_6/ptodata/1/pna/US6023B COMB.seq:*
- 94: /cgn2_6/ptodata/1/pna/US6024 COMB.seq:*
- 95: /cgn2_6/ptodata/1/pna/US6025 COMB.seq:*
- 96: /cgn2_6/ptodata/1/pna/US6026 COMB.seq:*
- 97: /cgn2_6/ptodata/1/pna/US6027 COMB.seq:*
- 98: /cgn2_6/ptodata/1/pna/US6028 COMB.seq:*
- 99: /cgn2_6/ptodata/1/pna/US6029 COMB.seq:*
- 100: /cgn2_6/ptodata/1/pna/US6030 COMB.seq:*
- 101: /cgn2_6/ptodata/1/pna/US6031 COMB.seq:*
- 102: /cgn2_6/ptodata/1/pna/US6032 COMB.seq:*
- 103: /cgn2_6/ptodata/1/pna/US6033 COMB.seq:*
- 104: /cgn2_6/ptodata/1/pna/US6034 COMB.seq:*
- 105: /cgn2_6/ptodata/1/pna/US6035 COMB.seq:*
- 106: /cgn2_6/ptodata/1/pna/US6036 COMB.seq:*
- 107: /cgn2_6/ptodata/1/pna/US6037 COMB.seq:*
- 108: /cgn2_6/ptodata/1/pna/US6038 COMB.seq:*
- 109: /cgn2_6/ptodata/1/pna/US6039 COMB.seq:*
- 110: /cgn2_6/ptodata/1/pna/US6040 COMB.seq:*
- 111: /cgn2_6/ptodata/1/pna/US6041 COMB.seq:*
- 112: /cgn2_6/ptodata/1/pna/US6042 COMB.seq:*
- 113: /cgn2_6/ptodata/1/pna/US6043 COMB.seq:*
- 114: /cgn2_6/ptodata/1/pna/US6044 COMB.seq:*
- 115: /cgn2_6/ptodata/1/pna/US6045 COMB.seq:*
- 116: /cgn2_6/ptodata/1/pna/US6046 COMB.seq:*

117: /cgn2_6/ptodata/1/pna/US6047_COMB.seq.*
 118: /cgn2_6/ptodata/1/pna/US6048_COMB.seq.*
 119: /cgn2_6/ptodata/1/pna/US6049_COMB.seq.*
 120: /cgn2_6/ptodata/1/pna/US6050_COMB.seq.*
 121: /cgn2_6/ptodata/1/pna/US6051_COMB.seq.*
 122: /cgn2_6/ptodata/1/pna/US6052_COMB.seq.*
 123: /cgn2_6/ptodata/1/pna/US6053_COMB.seq.*
 124: /cgn2_6/ptodata/1/pna/US6054_COMB.seq.*
 125: /cgn2_6/ptodata/1/pna/US6055_COMB.seq.*
 126: /cgn2_6/ptodata/1/pna/US6056_COMB.seq.*
 127: /cgn2_6/ptodata/1/pna/US6057_COMB.seq.*
 128: /cgn2_6/ptodata/1/pna/US6058_COMB.seq.*
 129: /cgn2_6/ptodata/1/pna/US6059_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1053	100.0	1053	58	US-10-698-228-12	Sequence 12, Appli
2	1053	100.0	1081	58	US-10-698-228-4	Sequence 4, Appli
3	1051.4	99.8	1053	58	US-10-698-228-2	Sequence 2, Appli
4	1051.4	99.8	2596	34	US-09-787-491B-28	Sequence 28, Appli
5	1051.4	99.8	2596	81	US-60-128-660-20	Sequence 20, Appli
6	1051.4	99.8	3397	49	US-10-170-235-20856	Sequence 20856, A
7	1051.4	99.8	4955	1	PCT-US02-18947-71	Sequence 71, Appli
8	1051.4	99.8	4955	49	US-10-172-118-71	Sequence 71, Appli
9	1051.4	99.8	4955	49	US-10-342-887-71	Sequence 71, Appli
10	1051.4	99.8	4955	52	US-10-698-228-3	Sequence 3, Appli
11	1051.4	99.8	1849	57	US-10-990-328-49	Sequence 49, Appli
12	1051.4	99.8	1849	68	US-10-698-228-3	Sequence 20, Appli
13	1051	99.8	3397	114	US-60-449-629-12	Sequence 12, Appli
14	1051	99.8	1601	28	US-09-629-469A-12810	Sequence 12810, A
15	1048.2	99.5	1601	65	US-10-917-503-12810	Sequence 266, App
16	1048.2	99.5	2597	97	US-60-278-258-266	Sequence 6398, App
17	1040.4	98.8	5609	26	US-09-584-852-6398	Sequence 870, App
18	987.4	83.1	1140	48	US-10-144-771-870	Sequence 870, App
19	878.4	83.4	1140	106	US-60-360-207-870	Sequence 870, App
20	878.4	83.4	1140	106	PCT-US03-13853-8394	Sequence 8394, App
21	850.6	61.8	706	56	US-10-513-369-8394	Sequence 8394, App
22	850.6	61.8	706	107	US-60-377-240-8394	Sequence 8394, App
23	850.6	61.8	706	107	US-09-658-659-9	Sequence 9, Appli
24	850.6	55.1	2500	29	US-09-658-659B-9	Sequence 9, Appli
25	850.6	55.1	2500	29	PCT-US01-04926A-514	Sequence 514, App
26	850.6	55.1	2482	1	US-09-515-126-13666	Sequence 13666, A
27	850.6	55.1	2482	23	US-09-577-408-13666	Sequence 514, App
28	850.6	55.1	2482	26	US-10-220-335-514	Sequence 514, App
29	850.6	55.1	2482	51	US-10-290-752-514	Sequence 4, Appli
30	850.6	55.1	2500	1	PCT-US02-10824-4	Sequence 724, App
31	850.6	55.1	2500	42	US-09-954-456-724	Sequence 1169, App
32	850.6	55.1	2500	42	US-09-954-456-1827	Sequence 1827, App
33	850.6	55.1	2500	42	US-09-954-456-1827	Sequence 1370, App
34	850.6	55.1	2500	58	US-10-641-643-1370	Sequence 458, App
35	850.6	55.1	2500	62	US-10-733-878-458	Sequence 713, App
36	850.6	55.1	2500	62	US-10-756-149-713	Sequence 3751, App
37	850.6	55.1	2500	63	US-10-843-641A-3751	Sequence 4196, App
38	850.6	55.1	2500	63	US-10-843-641A-4196	Sequence 4854, App
39	850.6	55.1	2500	66	US-10-940-774-145	Sequence 145, App
40	850.6	55.1	2500	66	US-11-000-688-1211	Sequence 1211, App
41	850.6	55.1	2500	126	US-60-568-073-328	Sequence 328, App
42	850.6	55.1	1649	2	PCT-US04-09289-583	Sequence 583, App
43	850.6	55.1	1649	1	PCT-US00-05882-505	Sequence 505, App
44	578.4	54.9	1899	1		
45	578.4	54.9				

ALIGNMENTS

RESULT 1
 US-10-698-228-12
 ; Sequence 12, Application US/10698228
 ; GENERAL INFORMATION:
 ; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
 ; TITLE OF INVENTION: New Protein and its DNA
 ; FILE REFERENCE: 2619WOOP
 ; CURRENT APPLICATION NUMBER: US/10/698,228
 ; CURRENT FILING DATE: 2003-10-30
 ; PRIOR APPLICATION NUMBER: US/10/019,733
 ; PRIOR FILING DATE: 2001-12-28
 ; PRIOR APPLICATION NUMBER: JP 11-181131
 ; PRIOR FILING DATE: 1999-06-28
 ; PRIOR APPLICATION NUMBER: JP 11-192391
 ; PRIOR FILING DATE: 1999-07-06
 ; PRIOR APPLICATION NUMBER: JP 2000-017770
 ; PRIOR FILING DATE: 2000-01-21
 ; NUMBER OF SEQ ID NOS: 14
 ; SEQ ID NO 12
 ; LENGTH: 1053
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION:
 US-10-698-228-12
 Query Match 100.0%; Score 1053; DB 58; Length 1053;
 Best local Similarity 100.0%; Pred. No. 6.1e-276;
 Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGGCGACCCGGAAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
 DB 1 ATGGGCGACCCGGAAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
 QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTGC 120
 DB 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTGC 120
 QY 121 CGGTTTGTATCTTTCATTCAGTACCTGATATTTGGAAATGTATAACAGGCACAG 180
 DB 121 CGGTTTGTATCTTTCATTCAGTACCTGATATTTGGAAATGTATAACAGGCACAG 180
 QY 181 GCTTCTCTTGGACAGAGAGGTTGACTTATCAAGAGGATCTCCCTCACTGGAAACAAG 240
 DB 181 GCTTCTCTTGGACAGAGAGGTTGACTTATCAAGAGGATCTCCCTCACTGGAAACAAG 240
 QY 241 CTTAAAGCAGATGAGAGTACTTTCATCTCTCACATCTTTAGCCTTTTTCAGCCAGTGAT 300
 DB 241 CTTAAAGCAGATGAGAGTACTTTCATCTCTCACATCTTTAGCCTTTTTCAGCCAGTGAT 300
 QY 301 GGAATTGTAAATGAAATTTGGTGGAGGCTTTAGTCAGGAGGTCAGGTTCCAGAGGT 360
 DB 301 GGAATTGTAAATGAAATTTGGTGGAGGCTTTAGTCAGGAGGTCAGGTTCCAGAGGT 360
 QY 361 CGCTGTTTCTATGCTTTCAAAATTTCTCATCAGAGATGTTCACTCAGAGATGTTACAGTTTG 420
 DB 361 CGCTGTTTCTATGCTTTCAAAATTTCTCATCAGAGATGTTCACTCAGAGATGTTACAGTTTG 420
 QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
 DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
 QY 481 ACCATGCCCTATGTTAAGAAAAAGAGATTGGGCTTGCATGATGATGAGATGAGAAAA 540
 DB 481 ACCATGCCCTATGTTAAGAAAAAGAGATTGGGCTTGCATGATGATGAGATGAGAAAA 540
 QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTCTCTAGAGAGTTTCTTCTCAGGA 600
 DB 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTCTCTAGAGAGTTTCTTCTCAGGA 600
 QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
 DB 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
 QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
 DB 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660

QY 661 AATGAATCATCAGCAGAGATGAAGACCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
DB 661 AATGAATCATCAGCAGAGATGAAGACCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
QY 721 TACTTAGTAATAAGCCTTCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
DB 721 TACTTAGTAATAAGCCTTCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
QY 781 ATTGAGCAGAGTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTCATT 840
DB 781 ATTGAGCAGAGTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTCATT 840
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGTGAACTTGTCA 900
DB 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGTGAACTTGTCA 900
QY 901 AAGGTTTTTCAGGCGAGAAATCCTTTGATTTTATGAGAAACATTTCTTTAGAGGAAA 960
DB 901 AAGGTTTTTCAGGCGAGAAATCCTTTGATTTTATGAGAAACATTTCTTTAGAGGAAA 960
QY 961 ACAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
DB 961 ACAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
QY 1021 ACAGATAACGCTCTTACCTTTGGATGCAGATTTT 1053
DB 1021 ACAGATAACGCTCTTACCTTTGGATGCAGATTTT 1053

RESULT 2

US-10-698-228-4
; Sequence 4, Application US/10698228
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-4

Query Match 100.0%; Score 1053; DB 58; Length 1081;
Best Local Similarity 100.0%; Pred. No. 6.2e-276;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGGACCCGGAAGGCGGAGCGGCTGGATCAGGATCAGAGATCATTTCA 60
DB 20 ATGGGGGACCCGGAAGGCGGAGCGGCTGGATCAGGATCAGAGATCATTTCA 79
QY 61 GACACCAAGCAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 80 GACACCAAGCAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 139
QY 121 CGGTTTGTATCTTTCCAAATCCAGTACCTGTATTTGGAAATGTATAAACAGGCACAG 180
DB 140 CGGTTTGTATCTTTCCAAATCCAGTACCTGTATTTGGAAATGTATAAACAGGCACAG 199
QY 181 GCTTCCTCTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAAG 240

DB 200 GCTTCTCTTGGACAGCAGAAAGAGTTGACTTATCAAAAGGATCTCCCTCACTGGAACAAG 259
QY 241 CTTTAAAGCAGATGAGAGTACTTCTCATCATCTTACATCTTACGCTTTTTCAGAGCCAGTAT 300
DB 260 CTTTAAAGCAGATGAGAGTACTTCTCATCATCTTACATCTTACGCTTTTTCAGAGCCAGTAT 319
QY 301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGAGGTCAGGTTCCAGAGGCT 360
DB 320 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGAGGTCAGGTTCCAGAGGCT 379
QY 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGATGTTCTCATCAGAGATGTAAGTTTG 420
DB 380 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGATGTTCTCATCAGAGATGTAAGTTTG 439
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 480
DB 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 499
QY 481 ACCATGCCCTATCTTAAGAAAGGAGATTTGGSCCTTTCGATGATAGCAGATAGAA 540
DB 500 ACCATGCCCTATCTTAAGAAAGGAGATTTGGSCCTTTCGATGATAGCAGATAGAA 559
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
DB 560 TCTACTTTTGGGGAAGAGTGGTGGCCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 619
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGGTCCTTATGCCAGAGTCACTTTTTC 660
DB 620 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGGTCCTTATGCCAGAGTCACTTTTTC 679
QY 661 AATGAATCATCAGCAGAGATGAAGACCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
DB 680 AATGAATCATCAGCAGAGATGAAGACCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 739
QY 721 TACTTAGTAATAAGCCTTCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
DB 740 TACTTAGTAATAAGCCTTCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAA 799
QY 781 ATTGAGCAGAGTTTAAACAGAGCCTTCCAGTTTGGCCTCATTTGGAATGAATTCAT 840
DB 800 ATTGAGCAGAGTTTAAACAGAGCCTTCCAGTTTGGCCTCATTTGGAATGAATTCAT 859
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGTGGAACCTTGGATTC 900
DB 860 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGTGGAACCTTGGATTC 919
QY 901 AAGGTTTTTCAGGCGAGAAATCCTTTGATTTTATGAAACATTTCTTTTAAAGAGGAAA 960
DB 920 AAGGTTTTTCAGGCGAGAAATCCTTTGATTTTATGAAACATTTCTTTTAAAGAGGAAA 979
QY 961 ACAATTTCTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
DB 980 ACAATTTCTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1039
QY 1021 ACAGATAACGCTCTTACCTTTGGATGCAGATTTT 1053
DB 1040 ACAGATAACGCTCTTACCTTTGGATGCAGATTTT 1072

RESULT 3

US-10-698-228-2
; Sequence 2, Application US/10698228
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28

;; PRIOR APPLICATION NUMBER: JP 11-192391
;; PRIOR FILING DATE: 1999-07-06
;; PRIOR APPLICATION NUMBER: JP 2000-017770
;; PRIOR FILING DATE: 2000-01-21
;; NUMBER OF SEQ ID NOS: 14
;; SEQ ID NO 2
;; LENGTH: 1053
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION:
US-10-698-228-2

Query Match 99.8%; Score 1051.4; DB 58; Length 1053;
Best Local Similarity 99.9%; Pred. No. 1.7e-275;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGGACCCGGAAGCGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 60
DB 1 ATGGCGGACCCGGAAGCGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 60

QY 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAGAGTTCTCG 120
DB 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAGAGTTCTCG 120

QY 121 CGTTTGTGTCATCTTCCCAATCCAGTACCCTGATATTTGAAAATGTATAAACAGGCACAG 180
DB 121 CGTTTGTGTCATCTTCCCAATCCAGTACCCTGATATTTGAAAATGTATAAACAGGCACAG 180

QY 181 GCTTCCTTCTGACAGCAGAGAGGTGACTTATCAAGAGTCTCCCTCACTGGAAACAG 240
DB 181 GCTTCCTTCTGACAGCAGAGAGGTGACTTATCAAGAGTCTCCCTCACTGGAAACAG 240

QY 241 CTTAAGCAGATGAGAGTACTTCTCTCACAATCTAGCCCTTTTTCAGCCAGTGTAT 300
DB 241 CTTAAGCAGATGAGAGTACTTCTCTCACAATCTAGCCCTTTTTCAGCCAGTGTAT 300

QY 301 GGAATTTGAATGAATAATTTGGTGAGCGCTTTAGTCAGAGGTGCGAGGTCCAGAGGCT 360
DB 301 GGAATTTGAATGAATAATTTGGTGAGCGCTTTAGTCAGAGGTGCGAGGTCCAGAGGCT 360

QY 361 CGCTGTTCTATGGCTTTCAAAATCTCATCGAAGTGTTCCTCAGAGATGTACAGTTTG 420
DB 361 CGCTGTTCTATGGCTTTCAAAATCTCATCGAAGTGTTCCTCAGAGATGTACAGTTTG 420

QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA 480
DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA 480

QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTCGGCTTCGGATGGATAGCAGATAGAAA 540
DB 481 ACCATGCCCTATGTTAAGAAAAGCAGATTCGGCTTCGGATGGATAGCAGATAGAAA 540

QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTGTAGAAAGAGTCTTCTTCTCAGGA 600
DB 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTGTAGAAAGAGTCTTCTTCTCAGGA 600

QY 601 TCTTTTGTGCTATATCTGGCTTAAAGAAAGAGAGTCTTATGCAAGACTCTTTTTC 660
DB 601 TCTTTTGTGCTATATCTGGCTTAAAGAAAGAGAGTCTTATGCAAGACTCTTTTTC 660

QY 661 AATGAACATCAGCAGAGATCAAGGATTCCTGCTGACTTTGCTTGCCTGATGTTCCAA 720
DB 661 AATGAACATCAGCAGAGATCAAGGATTCCTGCTGACTTTGCTTGCCTGATGTTCCAA 720

QY 721 TACTTAGTAAATAAGCTTCAAGAAAAGGCTCAGGAGATCATTTGCTGCTGTCAAA 780
DB 721 TACTTAGTAAATAAGCTTCAAGAAAAGGCTCAGGAGATCATTTGCTGCTGTCAAA 780

QY 781 ATTGACGAGGAGTTTTTAAACAGAGCTTTGCCAGTTGGCTCATTTGGAATGAATTGCA 840
DB 781 ATTGACGAGGAGTTTTTAAACAGAGCTTTGCCAGTTGGCTCATTTGGAATGAATTGCA 840

RESULT 4

US-09-787-491B-28
; Sequence 28, Application US/09787491B
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; TANG, Y. Tom;
; APPLICANT: CORLEY, Neil C.; GUEGLER, Karl J.;
; APPLICANT: GORGONE, Gina A.; PATTERSON, Chandra;
; APPLICANT: HILLMAN, Jennifer L.; BAUGHN, Mariah R.;
; APPLICANT: LAL, Preeti; AZIMZAI, Yalda;
; APPLICANT: YUE, Henry; YANG, Junning
; TITLE OF INVENTION: RNA-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0600 USN
; CURRENT APPLICATION NUMBER: US/09/787,491B
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: PCT/US99/21688
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US 60/128,660
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: US 60/069,391
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: US 60/183,025
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: US 60/155,246
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: US 09/158,720
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PERL Program
; SEQ ID NO 28
; LENGTH: 2596
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 2604449CB1
US-09-787-491B-28

Query Match 99.8%; Score 1051.4; DB 34; Length 2596;
Best Local Similarity 99.9%; Pred. No. 2.4e-275;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 60
DB 56 ATGGGCGACCCGGAAGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 115

QY 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAGAGTTCTCG 120
DB 116 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAGAGTTCTCG 175

QY 121 CGTTTGTGTCATCTTCCCAATCCAGTACCCTGATATTTGAAAATGTATAAACAGGCACAG 180
DB 176 CGTTTGTGTCATCTTCCCAATCCAGTACCCTGATATTTGAAAATGTATAAACAGGCACAG 235

QY 181 GCTTCCTTCTGACAGCAGAGAGGTGACTTATCAAGAGTCTCCCTCACTGGAAACAG 240
DB 236 GCTTCCTTCTGACAGCAGAGAGGTGACTTATCAAGAGTCTCCCTCACTGGAAACAG 295

```
QY 241 CTTAAAGCAGATGAGAGTACTTCACTCTCACTCTTACCTTTTTCAGCCAGTGAT 300
Db 296 CTTAAAGCAGATGAGAGTACTTCACTCTCACTCTTACCTTTTTCAGCCAGTGAT 355
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGAGGCT 360
Db 356 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGAGGCT 415
QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAGTGTTCACATCAGAGTGTACGTTTG 420
Db 416 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAGTGTTCACATCAGAGTGTACGTTTG 475
QY 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTTGAA 480
Db 476 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTTGAA 535
QY 481 ACCATGCCCTATGTTAAGAAAGAGAGATGAGAGTGTTCAGATGAGATAGAGAA 540
Db 536 ACCATGCCCTATGTTAAGAAAGAGAGATGAGAGTGTTCAGATGAGATAGAGAA 595
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGAGTGTTCCTCTCAGGA 600
Db 596 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGAGTGTTCCTCTCAGGA 655
QY 601 TCTTTTGGCTGTATATTTCTGGCTTAAAGAGAGAGAGTCTTATGCGCAGACTCACCTTTTCC 660
Db 656 TCTTTTGGCTGTATATTTCTGGCTTAAAGAGAGAGTCTTATGCGCAGACTCACCTTTTCC 715
QY 661 AATGAACTCATCAGCAGATGAGAGTCTTCACTGTGACTTTTGGCTTGCCTGATGTTCCAA 720
Db 716 AATGAACTCATCAGCAGATGAGAGTCTTCACTGTGACTTTTGGCTTGCCTGATGTTCCAA 775
QY 721 TACTTAGTAAATAGCCTTCAGAGAAAGGCTCAGGAGATCAATGTTGATGCTGTCAA 780
Db 776 TACTTAGTAAATAGCCTTCAGAGAAAGGCTCAGGAGATCAATGTTGATGCTGTCAA 835
QY 781 ATTGAGCAGAGTGTTCACAGAGCCTTGCAGAGTGTGGCTTCAATGGAATGCAATTT 840
Db 836 ATTGAGCAGAGTGTTCACAGAGCCTTGCAGAGTGTGGCTTCAATGGAATGCAATTT 895
QY 841 TTGATGAAACAGTACATTTAGTGTGTAGCTGACAGATTTACTTGTGGAACCTGGATTTCTCA 900
Db 896 TTGATGAAACAGTACATTTAGTGTGTAGCTGACAGATTTACTTGTGGAACCTGGATTTCTCA 955
QY 901 AAGGTTTTTTCAGCAGAAATCCTTTTGAATTTATGGAAGAAATTTCTTTAGAGGAA 960
Db 956 AAGGTTTTTTCAGCAGAAATCCTTTTGAATTTATGGAAGAAATTTCTTTAGAGGAA 1015
QY 961 ACAATTTTCTTGGAGAAAGAGTTCAGAGTATCAGCGTTTTCAGATTTAGCAGAAACC 1020
Db 1016 ACAATTTTCTTGGAGAAAGAGTTCAGAGTATCAGCGTTTTCAGATTTAGCAGAAACC 1075
QY 1021 ACAGATAAGCTCTTCACTTGGATGAGAGTTT 1053
Db 1076 ACAGATAAGCTCTTCACTTGGATGAGAGTTT 1108
```

RESULT 5

```
US-60-128-660-20
; Sequence 20, Application US/60128660
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Tang, Y. Tom
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Yang, Junming
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: RNA-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0682 P
; CURRENT APPLICATION NUMBER: US/60128,660
; CURRENT FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 26
```

```
; SOFTWARE: PERL Program
; SEQ ID NO 20
; LENGTH: 2596
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 2604449
US-60-128-660-20
```

```
Query Match 99.8%; Score 1051.4; DB 81; Length 2596;
Best Local Similarity 99.9%; Pred. No. 2.4e-275;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 ATGGGCGACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 56 ATGGGCGACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 115
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 116 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 175
QY 121 CGGTTTGTGATCTTTTCAATTCAGTACCTGATATTTGGAAATATGATAAAGCAGCACAG 180
Db 176 CGGTTTGTGATCTTTTCAATTCAGTACCTGATATTTGGAAATATGATAAAGCAGCACAG 235
QY 181 GCTTCTCTTCTGGACAGCAGAGAGGTTGACTTATCAAGAGGATCTCCCTCACTGGAACAAG 240
Db 236 GCTTCTCTTCTGGACAGCAGAGAGGTTGACTTATCAAGAGGATCTCCCTCACTGGAACAAG 295
QY 241 CTTAAAGCAGATGAGAGTACTTCACTCTCACTCTTAGCCTTTTTCAGCCAGTGAT 300
Db 296 CTTAAAGCAGATGAGAGTACTTCACTCTCACTCTTAGCCTTTTTCAGCCAGTGAT 355
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGCTCCAGAGGCT 360
Db 356 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGCTCCAGAGGCT 415
QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCAGAGATGTTCACTCAGAGATGACAGTTTG 420
Db 416 CGCTGTTTCTATGGCTTTCAAAATTCATCAGAGATGTTCACTCAGAGATGACAGTTTG 475
QY 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTTCAA 480
Db 476 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTTGAA 535
QY 481 ACCATGCCCTATGTTAAGAAAGAGAGTGTGGCCTTTCGATGAGATGAGAGTGTTCCTCAGGA 540
Db 536 ACCATGCCCTATGTTAAGAAAGAGAGTGTGGCCTTTCGATGAGATGAGAGTGTTCCTCAGGA 595
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGAGTGTTCCTCAGGA 600
Db 596 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGAGTGTTCCTCAGGA 655
QY 601 TCTTTTGGCTGTATATTTCTGGCTTAAAGAGAGAGTCTTATGCGCAGACTCACCTTTTCC 660
Db 656 TCTTTTGGCTGTATATTTCTGGCTTAAAGAGAGAGTCTTATGCGCAGACTCACCTTTTCC 715
QY 661 AATGAACTCATCAGCAGATGAGAGTCTTCACTGTGACTTTTGGCTTGCCTGATGTTCCAA 720
Db 716 AATGAACTCATCAGCAGATGAGAGTCTTCACTGTGACTTTTGGCTTGCCTGATGTTCCAA 775
QY 721 TACTTAGTAAATAGCCTTCAGAGAAAGGCTCAGGAGATCAATGTTGATGCTGTCAA 780
Db 776 TACTTAGTAAATAGCCTTCAGAGAAAGGCTCAGGAGATCAATGTTGATGCTGTCAA 835
QY 781 ATTGAGCAGAGTGTTCACAGAGCCTTGCAGAGTGTGGCTTCAATGGAATGCAATTT 840
Db 836 ATTGAGCAGAGTGTTCACAGAGCCTTGCAGAGTGTGGCTTCAATGGAATGCAATTT 895
QY 841 TTGATGAAACAGTACATTTAGTGTGTAGCTGACAGATTTACTTGTGGAACCTGGATTTCTCA 900
Db 896 TTGATGAAACAGTACATTTAGTGTGTAGCTGACAGATTTACTTGTGGAACCTGGATTTCTCA 955
```



```
QY 901 AAGGTTTTTCAGCAGAGAAATCCTTTTGATTTATGGAAGAAATTTCTTTAGAGGAAAA 960
Db 956 AAGGTTTTTCAGCAGAGAAATCCTTTTGATTTATGGAAGAAATTTCTTTAGAGGAAAA 1015
QY 961 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGAGAAACC 1020
Db 1016 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGAGAAACC 1075
QY 1021 ACAGATAAGCTCTTACCTTTGGATGAGATTTT 1053
Db 1076 ACAGATAAGCTCTTACCTTTGGATGAGATTTT 1108

RESULT 6
US-10-170-235-20856
; Sequence 20856, Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
; FILE REFERENCE: CL001380
; CURRENT APPLICATION NUMBER: US/10/170,235
; NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 20856
; LENGTH: 3397
; TYPE: DNA
; ORGANISM: HUMAN
US-10-170-235-20856

Query Match 99.8%; Score 1051.4; DB 49; Length 3397;
Best Local Similarity 99.9%; Pred. No. 2.7e-275;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGACCCCGAAGCGCGAAGCGCGCGCTGGATCAGAGATGAGATCATCTTCA 60
Db 501 ATGGCGACCCCGAAGCGCGAAGCGCGCGCTGGATCAGAGATGAGATCATCTTCA 560
QY 61 GACACCAAGCAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAAGTTCTCCG 120
Db 561 GACACCAAGCAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAAGTTCTCCG 620
QY 121 CGGTTTGTTCATCTTCCAAATCCAGTACCCTGTATTTTGGAAATGTATAAAGGACACAG 180
Db 621 CGGTTTGTTCATCTTCCAAATCCAGTACCCTGTATTTTGGAAATGTATAAAGGACACAG 580
QY 181 GCTTCCTCTTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
Db 681 GCTTCCTCTTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 740
QY 241 CTTAAGCGAGATGAGAAAGTACTTCTCTCAGATCTTAGCCCTTTTTCAGCGCAGTGAT 300
Db 741 CTTAAGCGAGATGAGAAAGTACTTCTCTCAGATCTTAGCCCTTTTTCAGCGCAGTGAT 800
QY 301 GGAATTTGTAATGAATAATTTTGGTGGAGCGCTTTTGTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Db 801 GGAATTTGTAATGAATAATTTTGGTGGAGCGCTTTTGTAGTCAGAGGTGAGGTTCCAGAGGCT 860
QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db 861 CGCTGTTTCTATGGCTTTCAAATTTCTATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 920
QY 421 CTGATAGACATTAATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGAA 480
Db 921 CTGATAGACATTAATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGAA 980
QY 481 ACCATGCCCTATTTAAGAAAGAGAGATTTGGCCCTTGGATGATGAGATGAGATGAGAA 540
Db 981 ACCATGCCCTATTTAAGAAAGAGAGATTTGGCCCTTGGATGATGAGATGAGATGAGAA 1040
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTAGAGAGAGTTTCTTCTCAGGA 600
Db 1041 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTAGAGAGAGTTTCTTCTCAGGA 1100
```

```
QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGAAAGAGAGGCTTATGCCAGGACTCACATTTTCC 660
Db 1101 TCTTTTGTGCTATATTTCTGGCTAAAGAAAGAGAGGCTTATGCCAGGACTCACATTTTCC 1160
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTTCACCTGTGCACTTTTGTGCTGCTGATTTCAA 720
Db 1161 AATGAACCTCATCAGCAGAGATGAAGGACTTTCACCTGTGCACTTTTGTGCTGCTGATTTCAA 1220
QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTGCTGTCAA 780
Db 1221 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTGCTGTCAA 1280
QY 781 ATTGACGAGGAGTTTTTAAACAGAAAGCCTTCCAGTTGGCCTCATTTGCAATTCGCAAT 840
Db 1281 ATTGACGAGGAGTTTTTAAACAGAAAGCCTTCCAGTTGGCCTCATTTGCAATTCGCAAT 1340
QY 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATTTTGGAAACATTTCTTTAGAAAGAAA 960
Db 1341 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATTTTGGAAACATTTCTTTAGAAAGAAA 1400
QY 901 AAGGTTTTTCAGCAGAGAAATCCTTTTGATTTTATGGAACATTTCTTTAGAAAGAAA 960
Db 1401 AAGGTTTTTCAGCAGAGAAATCCTTTTGATTTTATGGAACATTTCTTTAGAAAGAAA 1460
QY 961 ACAAAATTTCTTTCAGAAAGCAGTTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1020
Db 1461 ACAAAATTTCTTTCAGAAAGCAGTTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1520
QY 1021 ACAGATAAGCTCTTTCACCTTTGGATGAGATTTT 1053
Db 1521 ACAGATAAGCTCTTTCACCTTTGGATGAGATTTT 1553
```

RESULT 7

```
PCT-US02-18947-71
; Sequence 71, Application PC/TUS0218947
; GENERAL INFORMATION:
; APPLICANT: Rosetta Inpharmatics
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-228
; CURRENT APPLICATION NUMBER: PCT/US02/18947
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AB036063
; DATABASE ENTRY DATE: 2001-06-18
PCT-US02-18947-71
```

```
Query Match 99.8%; Score 1051.4; DB 1; Length 4955;
Best Local Similarity 99.9%; Pred. No. 3.2e-275;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 ATGGCGACCCCGAAGCGCGAAGCGCGCGCTGGATCAGAGATGAGATCATCTTCA 60
Db 245 ATGGCGACCCCGAAGCGCGAAGCGCGCGCTGGATCAGAGATGAGATCATCTTCA 304
QY 61 GACACCAAGCAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 305 GACACCAAGCAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAAGAGTTCTCGC 364
QY 121 CGGTTTGTCTATCTTCCAAATTCAGATCCCTGATTTTGGAAATGTATAAAGAGGACAG 180
Db 365 CGGTTTGTCTATCTTCCAAATTCAGATCCCTGATTTTGGAAATGTATAAAGAGGACAG 424
QY 181 GCTTCCTCTTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
```

Db 425 GCTTCCTCTGACAGCAGAGAGGTCGACTTATCAAGAGATCTCCCTCACTGGAACAAG 484
Qy 241 CTTAAGCAGATGAGAAGTACTTCACTCTCATCTCTAGCCCTTTTTCGAGCCAGTGAT 300
Db 485 CTTAAGCAGATGAGAAGTACTTCACTCTCATCTCTAGCCCTTTTTCGAGCCAGTGAT 544
Qy 301 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Db 545 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 604
Qy 361 CGCTGTTCTATGCGCTTTCAAAATCTCATCGAAGTGTTCACCTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTCTATGCGCTTTCAAAATCTCATCGAAGTGTTCACCTCAGAGATGTACAGTTTG 664
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTGAA 724
Qy 481 ACCATGCGCTTATGTAAGAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAAA 540
Db 725 ACCATGCGCTTATGTAAGAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAAA 784
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGTGTTCCTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGTGTTCCTCTCAGGA 844
Qy 601 TCTTTTGTCTATATTTCTGGCTTAAGAAAGAGAGTCTTATGCGAGGACTCACTTTTTC 660
Db 845 TCTTTTGTCTATATTTCTGGCTTAAGAAAGAGAGTCTTATGCGAGGACTCACTTTTTC 904
Qy 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
Qy 721 TACTTAGTAAATAAGCCTTTCAGAGAAAGGCTCAGGAGATCATTTGATGCTGTCAAA 780
Db 965 TACTTAGTAAATAAGCCTTTCAGAGAAAGGCTCAGGAGATCATTTGATGCTGTCAAA 1024
Qy 781 ATTGAGCGAGTGTTCCTTTCAGAGAAAGGCTTGCAGTTGGCTCATTTGGAATGAATTCATT 840
Db 1025 ATTGAGCGAGTGTTCCTTTCAGAGAAAGGCTTGCAGTTGGCTCATTTGGAATGAATTCATT 1084
Qy 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATTTCA 900
Db 1085 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATTTCA 1144
Qy 901 AAGGTTTTTTCAGGCAAGAAATCCTTTTGAATTTTATGGAACAATTTCTTTAGAGGAAAA 960
Db 1145 AAGGTTTTTTCAGGCAAGAAATCCTTTTGAATTTTATGGAACAATTTCTTTAGAGGAAAA 1204
Qy 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1020
Db 1205 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1264
Qy 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
Db 1265 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1297

RESULT 8

PCT-US04-01100-71
; Sequence 71, Application PC/TUS0401100
; GENERAL INFORMATION:
; APPLICANT: Rosetta Inpharmatics LLC
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-228
; CURRENT APPLICATION NUMBER: PCT/US04/01100
; CURRENT FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: 10/342,887
; PRIOR FILING DATE: 2003-01-15
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955

; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US04-01100-71

Query Match 99.8%; Score 1051.4; DB 2; Length 4955;
Best Local Similarity 99.9%; Pred. No. 3.2e-275;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGCGCACCGGAAAGCGCCGGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 245 ATGGCGCACCGGAAAGCGCCGGGCTGGATCAGGATGAGAGATCATCTTCA 304
Qy 61 GACACCAACCAAGCTGAAATAAAGTCAAACTCAAGAGCCACTCTCTAAGAAAGAGTTCTCC 120
Db 305 GACACCAACCAAGCTGAAATAAAGTCAAACTCAAGAGCCACTCTCTAAGAAAGAGTTCTCC 364
Qy 121 CGGTTTGTCTATCTTTTCCAATCCAGTACCTCATATTTGGAATAATGTATAAACAGGCACAG 180
Db 365 CGGTTTGTCTATCTTTTCCAATCCAGTACCTCATATTTGGAATAATGTATAAACAGGCACAG 424
Qy 181 GCTTCTCTTGGACAGCAGAGAGGTTGACTTATCAAAAGGATCTCCCTCATCTGGAACAAG 240
Db 425 GCTTCTCTTGGACAGCAGAGAGGTTGACTTATCAAAAGGATCTCCCTCATCTGGAACAAG 484
Qy 241 CTTAAAGCAGATGAGAAGTACTTCTCTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT 300
Db 485 CTTAAAGCAGATGAGAAGTACTTCTCTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT 544
Qy 301 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Db 545 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 604
Qy 361 CGCTGTTCTATGCGCTTTCAAAATCTCATCGAAGTGTTCACCTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTCTATGCGCTTTCAAAATCTCATCGAAGTGTTCACCTCAGAGATGTACAGTTTG 664
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTGAA 724
Qy 481 ACCATGCGCTTATGTAAGAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAAA 540
Db 725 ACCATGCGCTTATGTAAGAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAAA 784
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGTGTTCCTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGTGTTCCTCTCAGGA 844
Qy 601 TCTTTTGTCTATATTTCTGGCTTAAGAAAGGCTTATGTCAGGACTCATCTTTTTC 660
Db 845 TCTTTTGTCTATATTTCTGGCTTAAGAAAGAGAGTCTTATGTCAGGACTCATCTTTTTC 904
Qy 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
Qy 721 TACTTAGTAAATAAGCCTTTCAGAGAAAGGCTCAGGAGATCATTTGATGCTGTCAAA 780
Db 965 TACTTAGTAAATAAGCCTTTCAGAGAAAGGCTCAGGAGATCATTTGATGCTGTCAAA 1024
Qy 781 ATTGAGCAGAGTGTTCCTTTCAGAGAAAGGCTTGCAGTTGGCTCATTTGGAATGAATTCATT 840
Db 1025 ATTGAGCAGAGTGTTCCTTTCAGAGAAAGGCTTGCAGTTGGCTCATTTGGAATGAATTCATT 1084
Qy 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATTTCA 900
Db 1085 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATTTCA 1144
Qy 901 AAGGTTTTTTCAGGCAAGAAATCCTTTTGAATTTTATGGAACAATTTCTTTAGAGGAAAA 960
Db 1145 AAGGTTTTTTCAGGCAAGAAATCCTTTTGAATTTTATGGAACAATTTCTTTAGAGGAAAA 1204
Qy 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1020

```

Db      1205 ACAAAATTCCTTTTGAGAAACGAGTTTCCAGAGTATCAGCGTTTTCAGGTATTGGCAGAACC 1364A
QY      1021 ACAGATAAACGCTCTTCACCTTGGATGCAGATTTT 1053
Db      1265 ACAGATAAACGCTCTTCACCTTGGATGCAGATTTT 1297

RESULT 9
US-10-172-118-71
; Sequence 71L, Application US/10172118
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-17S-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; PRIORITY FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLIC INFORMATION:
; DATABASE ACCESSION NUMBER: AB036063
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-71
```

[illegible]

RESULT 10
US-10-342-887-71
; Sequence 71, Application US/10342887
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-71

Query Match	99.8%	Score 1051.4;	DB 52;	Length 4955;
Best Local Similarity	99.9%	Pred. No. 3.2e-275;		
Matches 1052;	Conservative	0;	Mismatches 1;	Indels 0; Gaps 0;

```
QY 1 ATGGCGACCCCGAAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db |||||
QY 245 ATGGCGACCCCGAAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
Db |||||
QY 61 GACACCAAGAAAGTAAATGAAGTCAATGAAGAGCCACTCTTAAGAAAGTTCCTGC 120
Db |||||
QY 305 GACACCAAGAAAGTAAATGAAGTCAATGAAGAGCCACTCTTAAGAAAGTTCCTGC 364
Db |||||
QY 121 CGGTTTGTTCATCTTCCCAATCAGTACCTGATATTTGGAAATGATATAACAGGCACAG 180
Db |||||
QY 365 CGGTTTGTTCATCTTCCCAATCAGTACCTGATATTTGGAAATGATATAACAGGCACAG 424
Db |||||
QY 181 GCTTCCTCTTGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGAAACAG 240
Db |||||
QY 425 GCTTCCTCTTGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGAAACAG 484
Db |||||
QY 241 CTTAAAGCAGATGAGAGTACCTCTCTCATCTTACCTTTAGCCCTTTTGCAGCCAGTAT 300
Db |||||
QY 485 CTTAAAGCAGATGAGAGTACCTCTCTCATCTTACCTTTAGCCCTTTTGCAGCCAGTAT 544
Db |||||
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGCT 360
Db |||||
QY 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGCT 604
Db |||||
QY 361 CGCTGTTCTATGCGCTTTCAAATTCATCGAGATGTTGAGGATCTCCCTCACTGGAAACAG 420
Db |||||
QY 605 CGCTGTTCTATGCGCTTTCAAATTCATCGAGATGTTGAGGATCTCCCTCACTGGAAACAG 664
Db |||||
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGCAATTTTATTAATGCAATTGAA 480
Db |||||
QY 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGCAATTTTATTAATGCAATTGAA 724
Db |||||
QY 481 ACCATGCCCTATGTTTAAAGAAAAAGCAGATTTGGGCCCTTGGCATGGATAGCAGATAGAAAA 540
Db |||||
QY 725 ACCATGCCCTATGTTTAAAGAAAAAGCAGATTTGGGCCCTTGGCATGGATAGCAGATAGAAAA 784
Db |||||
QY 541 TCTACTTTTGGGAAAGAGTGTGCGCTTTGCTGTGTAAGAGGAGTTCCTCTCAGGA 600
Db |||||
QY 785 TCTACTTTTGGGAAAGAGTGTGCGCTTTGCTGTGTAAGAGGAGTTCCTCTCAGGA 844
Db |||||
QY 601 TCTTTTGTCTATCTTCTGGCTAAAGAGGAGTTCCTCTCAGGA 844
Db |||||
QY 845 TCTTTTGTCTATCTTCTGGCTAAAGAGGAGTTCCTCTCAGGA 904
Db |||||
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db |||||
QY 905 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
Db |||||
QY 721 TACTTAGTAAATAGCCCTTCAGAAAGAGGTCAGGAGATCAATGTTGATGCTGTCAA 780
Db |||||
QY 965 TACTTAGTAAATAGCCCTTCAGAAAGAGGTCAGGAGATCAATGTTGATGCTGTCAA 1024
Db |||||
QY 781 ATTGAGCAGGAGTTCCTTAAAGAAAGCCTTGGCCTTCAATGGAATGAATGCAAT 840
Db |||||
QY 1025 ATTGAGCAGGAGTTCCTTAAAGAAAGCCTTGGCCTTCAATGGAATGAATGCAAT 1084
Db |||||
QY 841 TTGATGAACAGTACATGAGTTCCTGATGATGATGATGATGATGATGATGATGATGATGAT 900
Db |||||
QY 1085 TTGATGAACAGTACATGAGTTCCTGATGATGATGATGATGATGATGATGATGATGATGAT 1144
Db |||||
QY 901 AAGGTTTTTCAGGCAGAAAAATCTTTGATTTTATGGAATAAATCTTTTGAAGAAAA 960
Db |||||
QY 1145 AAGGTTTTTCAGGCAGAAAAATCTTTGATTTTATGGAATAAATCTTTTGAAGAAAA 1204
Db |||||
QY 961 ACAAAATTTCTTTGAAAGAGGTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAAACC 1020
Db |||||
QY 1205 ACAAAATTTCTTTGAAAGAGGTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAAACC 1264
Db |||||
QY 1021 ACAGATAACGCTTTCACCTTTGGATGATGATTTT 1053
Db |||||
QY 1265 ACAGATAACGCTTTCACCTTTGGATGATGATTTT 1297
Db |||||
```

```
RESULT 11
US-10-698-228-3
; Sequence 3, Application US/10698228
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 3
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-3

Query Match 99.8%; Score 1051.4; DB 58; Length 4955;
Best Local Similarity 99.9%; Pred. No. 3.2e-275;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGACCCCGAAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db |||||
QY 245 ATGGCGACCCCGAAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
Db |||||
QY 61 GACACCAAGAAAGTAAATGAAGTCAATGAAGAGCCACTCTTAAGAAAGTTCCTGC 120
Db |||||
QY 305 GACACCAAGAAAGTAAATGAAGTCAATGAAGAGCCACTCTTAAGAAAGTTCCTGC 364
Db |||||
QY 121 CGGTTTGTTCATCTTCCCAATCAGTACCTGATATTTGGAAATGATATAACAGGCACAG 180
Db |||||
QY 365 CGGTTTGTTCATCTTCCCAATCAGTACCTGATATTTGGAAATGATATAACAGGCACAG 424
Db |||||
QY 181 GCTTCCTCTTGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGAAACAG 240
Db |||||
QY 425 GCTTCCTCTTGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGAAACAG 484
Db |||||
QY 241 CTTAAAGCAGATGAGAGTACCTCTCTCATCTTACCTTTAGCCCTTTTGCAGCCAGTAT 300
Db |||||
QY 485 CTTAAAGCAGATGAGAGTACCTCTCTCATCTTACCTTTAGCCCTTTTGCAGCCAGTAT 544
Db |||||
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGCT 360
Db |||||
QY 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGCT 604
Db |||||
QY 361 CGCTGTTCTATGCGCTTTCAAATTCATCGAGATGTTGAGGATCTCCCTCACTGGAAACAG 420
Db |||||
QY 605 CGCTGTTCTATGCGCTTTCAAATTCATCGAGATGTTGAGGATCTCCCTCACTGGAAACAG 664
Db |||||
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGCAATTTTATTAATGCAATTGAA 480
Db |||||
QY 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGCAATTTTATTAATGCAATTGAA 724
Db |||||
QY 481 ACCATGCCCTATGTTTAAAGAAAAAGCAGATTTGGGCCCTTGGCATGGATAGCAGATAGAAAA 540
Db |||||
QY 725 ACCATGCCCTATGTTTAAAGAAAAAGCAGATTTGGGCCCTTGGCATGGATAGCAGATAGAAAA 784
Db |||||
QY 541 TCTACTTTTGGGAAAGAGTGTGCGCTTTGCTGTGTAAGAGGAGTTCCTCTCAGGA 600
Db |||||
QY 785 TCTACTTTTGGGAAAGAGTGTGCGCTTTGCTGTGTAAGAGGAGTTCCTCTCAGGA 844
Db |||||
QY 601 TCTTTTGTCTATCTTCTGGCTAAAGAGGAGTTCCTCTCAGGA 844
Db |||||
QY 845 TCTTTTGTCTATCTTCTGGCTAAAGAGGAGTTCCTCTCAGGA 904
Db |||||
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db |||||
QY 905 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
Db |||||
QY 721 TACTTAGTAAATAGCCCTTCAGAAAGAGGTCAGGAGATCAATGTTGATGCTGTCAA 780
Db |||||
QY 965 TACTTAGTAAATAGCCCTTCAGAAAGAGGTCAGGAGATCAATGTTGATGCTGTCAA 1024
Db |||||
QY 781 ATTGAGCAGGAGTTCCTTAAAGAAAGCCTTGGCCTTCAATGGAATGAATGCAAT 840
Db |||||
QY 1025 ATTGAGCAGGAGTTCCTTAAAGAAAGCCTTGGCCTTCAATGGAATGAATGCAAT 1084
Db |||||
QY 841 TTGATGAACAGTACATGAGTTCCTGATGATGATGATGATGATGATGATGATGATGATGAT 900
Db |||||
QY 1085 TTGATGAACAGTACATGAGTTCCTGATGATGATGATGATGATGATGATGATGATGATGAT 1144
Db |||||
QY 901 AAGGTTTTTCAGGCAGAAAAATCTTTGATTTTATGGAATAAATCTTTTGAAGAAAA 960
Db |||||
QY 1145 AAGGTTTTTCAGGCAGAAAAATCTTTGATTTTATGGAATAAATCTTTTGAAGAAAA 1204
Db |||||
QY 961 ACAAAATTTCTTTGAAAGAGGTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAAACC 1020
Db |||||
QY 1205 ACAAAATTTCTTTGAAAGAGGTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAAACC 1264
Db |||||
QY 1021 ACAGATAACGCTTTCACCTTTGGATGATGATTTT 1053
Db |||||
QY 1265 ACAGATAACGCTTTCACCTTTGGATGATGATTTT 1297
Db |||||
```

QY 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
DB |||||
QY 905 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
DB |||||
QY 721 TACTTAGTAAATAAGCTTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGTCAAA 780
DB |||||
QY 965 TACTTAGTAAATAAGCTTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGTCAAA 1024
DB |||||
QY 781 ATTGACGAGAGTTTAAACAGAGCTTGCAGAGTTGGCTCATTTGGAATGAATTCATTT 840
DB |||||
QY 1025 ATTGACGAGAGTTTAAACAGAGCTTGCAGAGTTGGCTCATTTGGAATGAATTCATTT 1084
DB |||||
QY 841 TTGATGAACAGTACATTCAGTTTGTAGCTGCAGAGTTTACTTGTGAACTTGCATTTCTCA 900
DB |||||
QY 1085 TTGATGAACAGTACATTCAGTTTGTAGCTGCAGAGTTTACTTGTGAACTTGCATTTCTCA 1144
DB |||||
QY 901 AAGGTTTTTCAGGAGAGAAATCCTTTTGATTTTATGAGAAACATTTCTTTAGAGGAAAA 960
DB |||||
QY 1145 AAGGTTTTTCAGGAGAGAAATCCTTTTGATTTTATGAGAAACATTTCTTTAGAGGAAAA 1204
DB |||||
QY 961 ACAAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1020
DB |||||
QY 1205 ACAAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1264
DB |||||
QY 1021 ACAGATAAGCTCTTCACCTTTGGATGCAGATTTT 1053
DB |||||
QY 1265 ACAGATAAGCTCTTCACCTTTGGATGCAGATTTT 1297
DB |||||

RESULT 12

US-10-990-328-49

; Sequence 49, Application US/10990328

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele

; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES

; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND

; TITLE OF INVENTION: USES THEREOF

; FILE REFERENCE: CL001495

; CURRENT APPLICATION NUMBER: US/10/990,328

; NUMBER OF SEQ ID NOS: 558824

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 49

; LENGTH: 1849

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-990-328-49

Query Match 99.8%; Score 1051; DB 67; Length 1849;

Best Local Similarity 99.8%; Pred. No. 2.7e-275;

Matches 1051; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGGCGGAAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
DB |||||
QY 85 ATGGGCGACCCGGAAGGCGGAAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 144
DB |||||
QY 61 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB |||||
QY 145 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 204
DB |||||
QY 121 CGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAATGTTATTAACAGGCACAG 180
DB |||||
QY 205 CGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAATGTTATTAACAGGCACAG 264
DB |||||
QY 181 GCTTCTCTTGACAGCAGAGAGGTTGACTTATCAAGATCTCCCTCACTGGAACAAG 240
DB |||||
QY 265 GCTTCTCTTGACAGCAGAGAGGTTGACTTATCAAGATCTCCCTCACTGGAACAAG 324
DB |||||
QY 241 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCATCTTACGCTTTTTCAGCGCAGTGAT 300
DB |||||
QY 325 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCATCTTACGCTTTTTCAGCGCAGTGAT 384
DB |||||

QY 301 GGAATGTTAAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
DB |||||
QY 385 GGAATGTTAAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 444
DB |||||
QY 361 CGCTGTTTCTATGGCTTTCAAATTTCAATCGAGAAATGTTCACTCAGAGATGTCAGATTTG 420
DB |||||
QY 445 CGCTGTTTCTATGGCTTTCAAATTTCAATCGAGAAATGTTCACTCAGAGATGTCAGATTTG 504
DB |||||
QY 421 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTAATGCAATTCAGAA 480
DB |||||
QY 505 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTAATGCAATTCAGAA 564
DB |||||
QY 481 ACCATGCCCTATGCTTAAGAAAAAGCAGATTTGGCCCTTGGCATGCATAGCAGATAGAAA 540
DB |||||
QY 565 ACCATGCCCTATGCTTAAGAAAAAGCAGATTTGGCCCTTGGCATGCATAGCAGATAGAAA 624
DB |||||
QY 541 TCTACTTTTGGGGAAGAGTGGCTTTGCTGCTGCTAGAGAGTGTTCCTTCAGGA 600
DB |||||
QY 625 TCTACTTTTGGGGAAGAGTGGCTTTGCTGCTGCTAGAGAGTGTTCCTTCAGGA 684
DB |||||
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGCTCTTATGCCAGCACTCACCTTTTCC 660
DB |||||
QY 685 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGCTCTTATGCCAGCACTCACCTTTTCC 744
DB |||||
QY 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
DB |||||
QY 745 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 804
DB |||||
QY 721 TACTTAGTAAATAAGCTTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGTCTCAA 780
DB |||||
QY 805 TACTTAGTAAATAAGCTTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGTCTCAA 864
DB |||||
QY 781 ATTGACGAGAGTTTAAACAGAGCTTGCAGAGTTGGCTCATTTGGAATGAATTCATTT 840
DB |||||
QY 865 ATTGACGAGAGTTTAAACAGAGCTTGCAGAGTTGGCTCATTTGGAATGAATTCATTT 924
DB |||||
QY 841 TTGATGAACAGTACATTCAGTTTGTAGCTGCAGAGATTTACTTGTGAACTTTGGATTTCTCA 900
DB |||||
QY 925 TTGATGAACAGTACATTCAGTTTGTAGCTGCAGAGATTTACTTGTGAACTTTGGATTTCTCA 984
DB |||||
QY 901 AAGGTTTTTCAGGAGAGAAATCCTTTTGATTTTATGGAACAAATTTCTTTAGAGGAAAA 960
DB |||||
QY 985 AAGGTTTTTCAGGAGAGAAATCCTTTTGATTTTATGGAACAAATTTCTTTAGAGGAAAA 1044
DB |||||
QY 961 ACAAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1020
DB |||||
QY 1045 ACAAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1104
DB |||||
QY 1021 ACAGATAAGCTCTTCACCTTTGGATGCAGATTTT 1053
DB |||||
QY 1105 ACAGATAAGCTCTTCACCTTTGGATGCAGATTTT 1137
DB |||||

RESULT 13

US-60-505-218-20

; Sequence 20, Application US/60505218

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; TITLE OF INVENTION: CANCER, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001482

; CURRENT APPLICATION NUMBER: US/60/505,218

; CURRENT FILING DATE: 2003-09-24

; NUMBER OF SEQ ID NOS: 22507

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 20

; LENGTH: 1849

; TYPE: DNA

; ORGANISM: Homo sapiens

US-60-505-218-20

Query Match 99.8%; Score 1051; DB 120; Length 1849;

Best Local Similarity 99.8%; Pred. No. 2.7e-275;

Matches 1051; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGCGGAGCGGCTGGATCAGATCAGAGATCATCTTCA 60
DB |||||
QY 85 ATGGGCGACCCGGAAGCGCGGAGCGGCTGGATCAGATCAGAGATCATCTTCA 144
DB |||||
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB |||||
QY 145 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 204
DB |||||
QY 121 CGTTTGTCTATCTTCCAAATCGATACCTCTGATATTTGGAAATGATTAACAGGCACAG 180
DB |||||
QY 205 CGTTTGTCTATCTTCCAAATCGATACCTCTGATATTTGGAAATGATTAACAGGCACAG 264
DB |||||
QY 181 GCTTCTCTCTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGACACAG 240
DB |||||
QY 265 GCTTCTCTCTGACAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGACACAG 324
DB |||||
QY 241 CTTAAAGCAGATGAGAAAGTACTTCTCTCATCTTACATCTTAGCCCTTTTTCGAGCCAGTGAT 300
DB |||||
QY 325 CTTAAAGCAGATGAGAAAGTACTTCTCTCATCTTACATCTTAGCCCTTTTTCGAGCCAGTGAT 384
DB |||||
QY 301 GGAATGTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGAGTGAGGTCAGAGGCT 360
DB |||||
QY 385 GGAATGTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGAGTGAGGTCAGAGGCT 444
DB |||||
QY 361 CGCTGTTCTATGCGCTTTTCAAAATCTCATCGAGATGTTCTCATCAGAGATGACAGTTTG 420
DB |||||
QY 445 CGCTGTTCTATGCGCTTTTCAAAATCTCATCGAGATGTTCTCATCAGAGATGACAGTTTG 504
DB |||||
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTTGAA 480
DB |||||
QY 505 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTTGAA 564
DB |||||
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGCGATGAGATAGAGATAGAAA 540
DB |||||
QY 565 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGCGATGAGATAGAGATAGAAA 624
DB |||||
QY 541 TCTACTTTTGGGAAAGAGTGTTGGCTTTGCTGCTGATAGAGAGTTTCTTCTCAGGA 600
DB |||||
QY 625 TCTACTTTTGGGAAAGAGTGTTGGCTTTGCTGCTGATAGAGAGTTTCTTCTCAGGA 684
DB |||||
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAGAAAGAGGTTCTTATGCCAGACTCACTTTTCC 660
DB |||||
QY 685 TCTTTTGTGCTATATTTCTGGCTTAAGAAAGAGAGGTTCTTATGCCAGACTCACTTTTCC 744
DB |||||
QY 661 AATGAACCTCATCAGCAGAGATCAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
DB |||||
QY 745 AATGAACCTCATCAGCAGAGATCAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 804
DB |||||
QY 721 TACTTAGTAAATTAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
DB |||||
QY 805 TACTTAGTAAATTAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 864
DB |||||
QY 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATTCATTT 840
DB |||||
QY 865 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATTCATTT 924
DB |||||
QY 841 TTGATGAACACGTACATTCAGTTTGTGCTGACAGATTACTTGTGGAACCTTGGATTTCTCA 900
DB |||||
QY 925 TTGATGAACACGTACATTCAGTTTGTGCTGACAGATTACTTGTGGAACCTTGGATTTCTCA 984
DB |||||
QY 901 AAGGTTTTTTCAGGCAGAAAAATCTTTTGAATTTTATGGAATAACATTTCTTTAGAGGAAAA 960
DB |||||
QY 985 AAGGTTTTTTCAGGCAGAAAAATCTTTTGAATTTTATGGAATAACATTTCTTTAGAGGAAAA 1044
DB |||||
QY 961 ACAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGAGAAACC 1020
DB |||||
QY 1045 ACAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGAGAAACC 1104
DB |||||
QY 1021 ACAGATAAGCTTCTCACCTTTGGATGACAGATTTT 1053
DB |||||
QY 1105 ACAGATAAGCTTCTCACCTTTGGATGACAGATTTT 1137
DB |||||

RESULT 14

US-60-449-629-12
; Sequence 12, Application US/60449629
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CANCER, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001449
; CURRENT APPLICATION NUMBER: US/60/449,629
; CURRENT FILING DATE: 2003-02-04
; NUMBER OF SEQ ID NOS: 2712
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 3397
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-60-449-629-12

Query Match 99.8%; Score 1051; DB 114; Length 3397;
Best Local Similarity 99.8%; Pred. No. 3.5e-275;
Matches 1051; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGCGGAGCGGCTGGATCAGATCAGAGATCATCTTCA 60
DB |||||
QY 501 ATGGGCGACCCGGAAGCGCGGAGCGGCTGGATCAGATCAGAGATCATCTTCA 560
DB |||||
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB |||||
QY 561 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 620
DB |||||
QY 121 CGTTTGTCTATCTTCCAAATCGATACCTCTGATATTTGGAAATGATTAACAGGCACAG 180
DB |||||
QY 621 CGTTTGTCTATCTTCCAAATCGATACCTCTGATATTTGGAAATGATTAACAGGCACAG 680
DB |||||
QY 181 GCTTCTCTCTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAAACAG 240
DB |||||
QY 681 GCTTCTCTCTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAAACAG 740
DB |||||
QY 241 CTTAAAGCAGATGAGAAAGTACTTCTCATCTTCAATCTTAGCCCTTTTTCGAGCCAGTGAT 300
DB |||||
QY 741 CTTAAAGCAGATGAGAAAGTACTTCTCATCTTCAATCTTAGCCCTTTTTCGAGCCAGTGAT 800
DB |||||
QY 301 GGAATGTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGAGTGAGGTCAGAGGCT 360
DB |||||
QY 801 GGAATGTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGAGTGAGGTCAGAGGCT 860
DB |||||
QY 361 CGCTGTTCTATGCGCTTTTCAAAATCTCATCGAGATGTTCTCATCAGAGATGTAAGTTTG 420
DB |||||
QY 861 CGCTGTTCTATGCGCTTTTCAAAATCTCATCGAGATGTTCTCATCAGAGATGTAAGTTTG 920
DB |||||
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTTGAA 480
DB |||||
QY 921 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTTGAA 980
DB |||||
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGCGATGAGATAGAGATAGAAA 540
DB |||||
QY 981 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGCGATGAGATAGAGATAGAAA 1040
DB |||||
QY 541 TCTACTTTTGGGAAAGAGTGTTGGCTTTGCTGCTGATAGAGAGTTTCTTCTCAGGA 600
DB |||||
QY 1041 TCTACTTTTGGGAAAGAGTGTTGGCTTTGCTGCTGATAGAGAGTTTCTTCTCAGGA 1100
DB |||||
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAGAAAGAGAGGTTCTTATGCCAGACTCACTTTTCC 660
DB |||||
QY 1101 TCTTTTGTGCTATATTTCTGGCTTAAGAAAGAGAGGTTCTTATGCCAGACTCACTTTTCC 1160
DB |||||
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGATGTTCCAA 720
DB |||||
QY 1161 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGATGTTCCAA 1220
DB |||||
QY 721 TACTTAGTAAATTAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
DB |||||

[illegible]

RESULT 15

US-09-623-469A-12810
 ; Sequence 12810, Application US/09629469A
 ; GENERAL INFORMATION:
 ; APPLICANT: OTA, TOSHIO
 ; APPLICANT: ISOGAI, TAKAO
 ; APPLICANT: NISHIKAWA, TETSUO
 ; APPLICANT: HAYASHI, KOJI
 ; APPLICANT: SAITO, KAOBU
 ; APPLICANT: YAMAMOTO, JUNICHI
 ; APPLICANT: ISHII, SHIZUKO
 ; APPLICANT: SUGIYAMA, TOMOYASU
 ; APPLICANT: WAKAMATSU, AI
 ; APPLICANT: NAGAI, KEIICHI
 ; APPLICANT: OTSUKI, TETSUJI
 ; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE

Query Match	99.5%;	Score 1048.2;	DB 28;	Length 1601;
Best Local Similarity	99.7%;	Pred. No. 1.5e-274;		

QY 1 ATGGGCGACCCGGAAAGGCCCGGGCTGGATCAGGATGAGAGATCATCTTCA 60

Search completed: October 30, 2005, 05:15:34

Job time : 3217.5 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2005, 00:02:51 ; Search time 1372.5 Seconds
(without alignments)
11030.942 Million cell updates/sec

Title: US-10-698-228-12
Perfect score: 1053
Sequence: 1 atggcgaccggaagggcc.....tcaccttgatgcagatttt 1053

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27723106 seqs, 7188968421 residues

Total number of hits satisfying chosen parameters: 55446212

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA New:*

- 1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:*
- 2: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq2:*
- 3: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:*
- 4: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq:*
- 5: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:*
- 6: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:*
- 7: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq1:*
- 8: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq2:*
- 9: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:*
- 10: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq10:*
- 11: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq11:*
- 12: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq2:*
- 13: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq3:*
- 14: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq4:*
- 15: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq5:*
- 16: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq6:*
- 17: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq7:*
- 18: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq8:*
- 19: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq9:*
- 20: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:*
- 21: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq2:*
- 22: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq3:*
- 23: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq4:*
- 24: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq5:*
- 25: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq6:*
- 26: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq7:*
- 27: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq8:*
- 28: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1051.4	99.8	1056	1	PCT-US05-09639-16
2	1051.4	99.8	1056	27	US-11-088-686-16
3	1051.4	99.8	4929	1	PCT-US05-22501-1545
4	1051.4	99.8	4955	1	PCT-US05-09639-44
5	1051.4	99.8	4955	1	PCT-US05-21650-13

6	1051.4	99.8	4955	27	US-11-088-686-44	Sequence 44, Appl
7	1051	99.8	1849	15	US-10-990-328A-49	Sequence 49, Appl
8	1049.8	99.7	1056	1	PCT-US05-09639-18	Sequence 18, Appl
9	1049.8	99.7	1056	27	US-11-088-686-18	Sequence 18, Appl
10	1048.2	99.5	1056	1	PCT-US05-09639-20	Sequence 20, Appl
11	1048.2	99.5	1056	27	US-11-088-686-20	Sequence 20, Appl
12	1046.6	99.4	1056	1	PCT-US05-09639-22	Sequence 22, Appl
13	1046.6	99.4	1056	1	PCT-US05-09639-24	Sequence 24, Appl
14	1046.6	99.4	1056	27	US-11-088-686-22	Sequence 22, Appl
15	1046.6	99.4	1056	27	US-11-088-686-24	Sequence 24, Appl
16	1045	99.2	1056	1	PCT-US05-09639-26	Sequence 26, Appl
17	1045	99.2	1056	27	US-11-088-686-26	Sequence 26, Appl
18	1045	99.2	1056	27	US-11-088-686-28	Sequence 28, Appl
19	1045	99.2	1056	27	US-11-088-686-42	Sequence 42, Appl
20	624	59.3	624	1	PCT-US05-09639-42	Sequence 42, Appl
21	624	59.3	624	27	US-11-088-686-42	Sequence 42, Appl
22	580	55.1	1170	1	PCT-US05-09639-2	Sequence 2, Appl
23	580	55.1	1170	27	US-11-088-686-2	Sequence 2, Appl
24	580	55.1	2482	11	US-10-302-689A-107255	Sequence 107255,
25	580	55.1	2500	1	PCT-US05-09639-43	Sequence 43, Appl
26	580	55.1	2500	1	PCT-US05-20435-12	Sequence 12, Appl
27	580	55.1	2500	1	PCT-US05-21650-12	Sequence 12, Appl
28	580	55.1	2500	1	PCT-US05-27243-78	Sequence 78, Appl
29	580	55.1	2500	14	US-10-960-414-154	Sequence 154, App
30	580	55.1	2500	15	US-10-940-774A-145	Sequence 145, App
31	580	55.1	2500	24	US-11-000-688A-1211	Sequence 1211, Ap
32	580	55.1	2500	24	US-11-150-888-12	Sequence 12, Appl
33	578.4	54.9	1170	1	PCT-US05-09639-4	Sequence 4, Appl
34	578.4	54.9	1170	27	US-11-088-686-4	Sequence 4, Appl
35	578.4	54.9	1649	9	US-10-553-098-583	Sequence 583, App
36	578.4	54.9	1649	13	US-10-955-054A-135	Sequence 135, App
37	578.4	54.9	2470	1	PCT-US05-13260-66	Sequence 66, Appl
38	578.4	54.9	2470	1	PCT-US05-20840-66	Sequence 66, Appl
39	578.4	54.9	2470	1	PCT-US05-22501-6487	Sequence 6487, Ap
40	578.4	54.9	2479	15	US-10-940-774A-2025	Sequence 2025, Ap
41	578.4	54.9	2909	15	US-10-990-328A-4852	Sequence 4852, Ap
42	578	54.9	3616	15	US-10-990-328A-4851	Sequence 4851, Ap
43	578	54.9	1170	1	PCT-US05-09639-6	Sequence 6, Appl
44	576.8	54.8	1170	1	PCT-US05-09639-8	Sequence 8, Appl
45	576.8	54.8	1170	1	PCT-US05-09639-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
PCT-US05-09639-16
; Sequence 16, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1056)
PCT-US05-09639-16

Query Match 99.8%; Score 1051.4; DB 1; Length 1056;
Best Local Similarity 99.9%; Pred. No. 3.3e-270;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGCGACCGGAAGCGCGGAGCGCGGCTGCATCAGATCAGATCATCTTCA 60
|||||

Db 1 ATGGGCGACCCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACCAACGAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAGAGTTCTCGC 120
Db 61 GACACCAACGAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAGAGTTCTCGC 120
QY 121 CGGTTTGTGTCATCTTCCCAATCCAGTACCTCTGATATTGGAAATGTATAAAGAGGACAG 180
Db 121 CGGTTTGTGTCATCTTCCCAATCCAGTACCTCTGATATTGGAAATGTATAAAGAGGACAG 180
QY 181 GCTTCCTCTGACAGCAGAGAGTTGACTTATCAAGATCTCCTCACTGGAAACAAG 240
Db 181 GCTTCCTCTGACAGCAGAGAGTTGACTTATCAAGATCTCCTCACTGGAAACAAG 240
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCAATCTTAGCCCTTTTGGAGCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCAATCTTAGCCCTTTTGGAGCAGTGAT 300
QY 301 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGGTGAGGTTCCAGAGGCT 360
Db 301 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGGTGAGGTTCCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACATTTACATCAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 480
Db 421 CTGATAGACATTTACATCAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 480
QY 481 ACCATGCGCCCTATGTTAAGAAAGGAGATGCGGCTTGGATGAGTAGCAGATAGAAAA 540
Db 481 ACCATGCGCCCTATGTTAAGAAAGGAGATGCGGCTTGGATGAGTAGCAGATAGAAAA 540
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAAGAGGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAAGAGGAGTTTCTTCTCAGGA 600
QY 601 TCTTTTGTGCTATATTCTGGCTTAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
Db 601 TCTTTTGTGCTATATTCTGGCTTAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
QY 661 AATGAACCTATCAGCAGATGAGGACTTCACTGTGCACTTTGCTGCTGATGTTCCAA 720
Db 661 AATGAACCTATCAGCAGATGAGGACTTCACTGTGCACTTTGCTGCTGATGTTCCAA 720
QY 721 TACTTAGTAAATAAGCCTTCAGAAGAAAGGCTCAGGAGATCATTTGATGCTGTCAA 780
Db 721 TACTTAGTAAATAAGCCTTCAGAAGAAAGGCTCAGGAGATCATTTGATGCTGTCAA 780
QY 781 ATTGAGCAGAGTTTAAACAGAGCTTGCAGTTGGCTCATTTGGAATGAATTCGATT 840
Db 781 ATTGAGCAGAGTTTAAACAGAGCTTGCAGTTGGCTCATTTGGAATGAATTCGATT 840
QY 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAATCTTGATTTCTCA 900
Db 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAATCTTGATTTCTCA 900
QY 901 AAGGTTTTTTCAGGCGAATAATCCTTTGATTTTATGGAATAATTTCTTTAGAGGAAAA 960
Db 901 AAGGTTTTTTCAGGCGAATAATCCTTTGATTTTATGGAATAATTTCTTTAGAGGAAAA 960
QY 961 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTTATGAGCAAAACC 1020
Db 961 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTTATGAGCAAAACC 1020
QY 1021 ACAGATAAGCTTTTACCTTGTGATGAGATTTT 1053
Db 1021 ACAGATAAGCTTTTACCTTGTGATGAGATTTT 1053

; Sequence 16, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; FILE OF INVENTION: DRUG SCREENING
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1056)
US-11-088-686-16

Query Match 99.8%; Score 1051.4; DB 27; Length 1056;
Best Local Similarity 99.9%; Pred. No. 3.3e-270;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGGCGACCCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACCAACGAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAGAGTTCTCGC 120
Db 61 GACACCAACGAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAGAGTTCTCGC 120
QY 121 CGGTTTGTGTCATCTTCCCAATCCAGTACCTCTGATATTGGAAATGTATAAAGAGGACAG 180
Db 121 CGGTTTGTGTCATCTTCCCAATCCAGTACCTCTGATATTGGAAATGTATAAAGAGGACAG 180
QY 181 GCTTCCTCTGACAGCAGAGAGTTGACTTATCAAGATCTCCTCACTGGAAACAAG 240
Db 181 GCTTCCTCTGACAGCAGAGAGTTGACTTATCAAGATCTCCTCACTGGAAACAAG 240
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCAATCTTAGCCCTTTTGGAGCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCAATCTTAGCCCTTTTGGAGCAGTGAT 300
QY 301 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGGTGAGGTTCCAGAGGCT 360
Db 301 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGGTGAGGTTCCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACATTTACATCAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 480
Db 421 CTGATAGACATTTACATCAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 480
QY 481 ACCATGCGCCCTATGTTAAGAAAGGAGATGCGGCTTGGATGAGTAGCAGATAGAAAA 540
Db 481 ACCATGCGCCCTATGTTAAGAAAGGAGATGCGGCTTGGATGAGTAGCAGATAGAAAA 540
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAAGAGGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAAGAGGAGTTTCTTCTCAGGA 600
QY 601 TCTTTTGTGCTATATTCTGGCTTAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
Db 601 TCTTTTGTGCTATATTCTGGCTTAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
QY 661 AATGAACCTATCAGCAGATGAGGACTTCACTGTGCACTTTGCTGCTGATGTTCCAA 720
Db 661 AATGAACCTATCAGCAGATGAGGACTTCACTGTGCACTTTGCTGCTGATGTTCCAA 720
QY 721 TACTTAGTAAATAAGCCTTCAGAAGAAAGGCTCAGGAGATCATTTGATGCTGTCAA 780

```
Db 721 TACTTAGTAATAAGCCTTCAGAAAGGAGGTTCAGGAGATCATTTGATGCTGTCAAA 780
QY 781 ATTGACGAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTCGATT 840
Db 781 ATTGACGAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTCGATT 840
QY 841 TTGATGAACACAGTACATTCAGTTGAGTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Db 841 TTGATGAACACAGTACATTCAGTTGAGTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
QY 901 AAGGTTTTTCAGGAGAGAAATCCTTTTGATTTTATGAGAAACATTTCTTTAGAGGAAAA 960
Db 901 AAGGTTTTTCAGGAGAGAAATCCTTTTGATTTTATGAGAAACATTTCTTTAGAGGAAAA 960
QY 961 ACAAAATTTCTTGGAGAACGATTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAC 1020
Db 961 ACAAAATTTCTTGGAGAACGATTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAC 1020
QY 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053

RESULT 3
PCT-US05-22501-1545
; Sequence 1545, Application PC/TUS0522501
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals
; TITLE OF INVENTION: Determining Cancer-Linked Gene and Therapeutic Targets Using
; FILE REFERENCE: Molecular Cytogenetic Methods
; CURRENT APPLICATION NUMBER: PCT/US05/22501
; CURRENT FILING DATE: 2005-07-07
; PRIOR APPLICATION NUMBER: 60/581,699
; PRIOR FILING DATE: 2004-06-23
; NUMBER OF SEQ ID NOS: 7840
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1545
; LENGTH: 4929
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: cDNA sequence
PCT-US05-22501-1545

Query Match 99.8%; Score 1051.4; DB 1; Length 4929;
Best Local Similarity 99.9%; Pred. No. 5.1e-270;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGGGACCCGGAAGCGCGGAGCGGCTGGATTCAGAGTGAAGATCATCTTCA 60
Db 245 ATGGGGGACCCGGAAGCGCGGAGCGGCTGGATTCAGAGTGAAGATCATCTTCA 304
QY 61 GACACCAAGGAAGTGAATAAGTCAATGAAGACCACTCTTAAGAAAGGTTCTCCG 120
Db 305 GACACCAAGGAAGTGAATAAGTCAATGAAGACCACTCTTAAGAAAGGTTCTCCG 364
QY 121 CGGTTTGTTCATCTTCCCAATCCAGTACCTGATATTGGAAATGTATATAACAGGCACAG 180
Db 365 CGGTTTGTTCATCTTCCCAATCCAGTACCTGATATTGGAAATGTATATAACAGGCACAG 424
QY 181 GCTTCCTTTCGACAGCAGAGAGGTTGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 240
Db 425 GCTTCCTTTCGACAGCAGAGAGGTTGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 484
QY 241 CTTAAGCAGATGAGAACTACTTCACTCTCATCTTACCTTTTAGCCCTTTTTCAGCCAGTGT 300
Db 485 CTTAAGCAGATGAGAACTACTTCACTCTCATCTTACCTTTTAGCCCTTTTTCAGCCAGTGT 544
QY 301 GGAATTTGTAATGAAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGGTTCCAGAGGCT 360
Db 545 GGAATTTGTAATGAAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGGTTCCAGAGGCT 604
```

```
QY 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAAATGCAATGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAAATGCAATGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCCCTTGCATGGATGAGCAGATAGAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCCCTTGCATGGATGAGCAGATAGAAA 784
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTTCTCTCTGTAGAAAGAGTTTCTTCTCAGCA 600
Db 785 TCTACTTTTGGGGAAGAGTGGTGGCTTTTCTCTCTGTAGAAAGAGTTTCTTCTCAGCA 844
QY 601 TCTTTTGTCTCTATTTCTTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACATTTTTC 660
Db 845 TCTTTTGTCTCTATTTCTTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACATTTTTC 904
QY 661 AATGAATCTCATCAGCAGAGATGAAGACTTCTCAGTGTGACTTTTGTCTTGCCTGATGTTCAA 720
Db 905 AATGAATCTCATCAGCAGAGATGAAGACTTCTCAGTGTGACTTTTGTCTTGCCTGATGTTCAA 964
QY 721 TACTTAGTAATAAGCCTTCAGAGAAAGGTCAGGAGAGATCATTTGATGCTGTCAAA 780
Db 965 TACTTAGTAATAAGCCTTCAGAGAAAGGTCAGGAGAGATCATTTGATGCTGTCAAA 1024
QY 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTGTGGCCCTCATTTGGAATGAATTCGATT 840
Db 1025 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTGTGGCCCTCATTTGGAATGAATTCGATT 1084
QY 841 TTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTTACTTTGGAACTTTGGATTTCTCA 900
Db 1085 TTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTTACTTTGGAACTTTGGATTTCTCA 1144
QY 901 AAGGTTTTTCAGGAGAGAAATCCTTTTGATTTTATGAAACATTTCTTTAGAGGAAAA 960
Db 1145 AAGGTTTTTCAGGAGAGAAATCCTTTTGATTTTATGAAACATTTCTTTAGAGGAAAA 1204
QY 961 ACAAAATTTCTTTCAGAGAACGATTTTCAGAGTATCAGCGTTTTCAGTGTATGCGCAGAAC 1020
Db 1205 ACAAAATTTCTTTCAGAGAACGATTTTCAGAGTATCAGCGTTTTCAGTGTATGCGCAGAAC 1264
QY 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
Db 1265 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1297
```

```
RESULT 4
PCT-US05-09639-44
; Sequence 44, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-44
```

```
Query Match 99.8%; Score 1051.4; DB 1; Length 4955;
Best Local Similarity 99.9%; Pred. No. 5.1e-270;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 ATGGCGACCCCGAAGCGCGGAGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 245 ATGGCGACCCCGAAGCGCGGAGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
QY 61 GACACCAAGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAGAGTTCTCGC 120
Db 305 GACACCAAGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAGAGTTCTCGC 364
QY 121 CGGTTTGTTCATCTTCCCAATCCAGTACCTGATATTTGAAAATGATATAAAGAGGACAG 180
Db 365 CGGTTTGTTCATCTTCCCAATCCAGTACCTGATATTTGAAAATGATATAAAGAGGACAG 424
QY 181 GCTTCTCTTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCAGGACAGCAAG 240
Db 425 GCTTCTCTTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCAGGACAGCAAG 484
QY 241 CTTAAAGCAGATGAGAGTACTTCACTCTCAATCTAGGCTTTTTCGAGCCAGTGAT 300
Db 485 CTTAAAGCAGATGAGAGTACTTCACTCTCAATCTAGGCTTTTTCGAGCCAGTGAT 544
QY 301 GGAATTTGTAATGAATAATTTGTGGAGCGCTTTAGTCAGGAGTGCGAGGCT 360
Db 545 GGAATTTGTAATGAATAATTTGTGGAGCGCTTTAGTCAGGAGTGCGAGGCT 604
QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTCGAGAATGTTCACTCAGAGATGTACAGTTG 420
Db 605 CGCTGTTTCTATGGCTTTCAAATTTCTCGAGAATGTTCACTCAGAGATGTACAGTTG 664
QY 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 480
Db 665 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 724
QY 481 ACCATGCCCTATCTTAAGAAAAGCAGATTTGGCCCTTGGATGATAGAGATAGAAA 540
Db 725 ACCATGCCCTATCTTAAGAAAAGCAGATTTGGCCCTTGGATGATAGAGATAGAAA 784
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCCCTTTGCTGTAGAAAGAGTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGGTGGCCCTTTGCTGTAGAAAGAGTTTCTTCTCAGGA 844
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGGCTTATTCAGGACTCATTCTTCC 660
Db 845 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGGCTTATTCAGGACTCATTCTTCC 904
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTCCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTCCAA 964
QY 721 TACTTAGTAATAAGCCTTCAGAAAGAGGCTCAGGAGAGATCATTTGTTGATGCTGTCAA 780
Db 964 TACTTAGTAATAAGCCTTCAGAAAGAGGCTCAGGAGAGATCATTTGTTGATGCTGTCAA 780
```

RESULT 5

```
PCT-US05-21650-13
; Sequence 13, Application PC/TUS0521650
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: MIGFS AS MODIFIERS OF THE IGF PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX05-017
; CURRENT APPLICATION NUMBER: PCT/US05/21650
; CURRENT FILING DATE: 2005-06-27
; PRIOR APPLICATION NUMBER: 60/581,689
; PRIOR FILING DATE: 2004-06-21
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-21650-13
```

```
Query Match 99.8%; Score 1051.4; DB 1; Length 4955;
Best Local Similarity 99.9%; Pred. No. 5.1e-270;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGACCCCGAAGCGCGGAGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 245 ATGGCGACCCCGAAGCGCGGAGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
QY 61 GACACCAAGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAGAGTTCTCGC 120
Db 305 GACACCAAGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAGAGTTCTCGC 364
QY 121 CGGTTTGTTCATCTTCCCAATCCAGTACCTGATATTTGAAAATGATATAAAGAGGACAG 180
Db 365 CGGTTTGTTCATCTTCCCAATCCAGTACCTGATATTTGAAAATGATATAAAGAGGACAG 424
QY 181 GCTTCTCTTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCAGGACAGCAAG 240
Db 425 GCTTCTCTTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCAGGACAGCAAG 484
QY 241 CTTAAAGCAGATGAGAGTACTTCACTCTCAATCTAGGCTTTTTCGAGCCAGTGAT 300
Db 485 CTTAAAGCAGATGAGAGTACTTCACTCTCAATCTAGGCTTTTTCGAGCCAGTGAT 544
QY 301 GGAATTTGTAATGAATAATTTGTGGAGCGCTTTAGTCAGGAGTGCGAGGCT 360
Db 545 GGAATTTGTAATGAATAATTTGTGGAGCGCTTTAGTCAGGAGTGCGAGGCT 604
QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTCGAGAATGTTCACTCAGAGATGTACAGTTG 420
Db 605 CGCTGTTTCTATGGCTTTCAAATTTCTCGAGAATGTTCACTCAGAGATGTACAGTTG 664
QY 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 480
Db 665 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 724
QY 481 ACCATGCCCTATCTTAAGAAAAGCAGATTTGGCCCTTGGATGATAGAGATAGAAA 540
Db 725 ACCATGCCCTATCTTAAGAAAAGCAGATTTGGCCCTTGGATGATAGAGATAGAAA 784
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCCCTTTGCTGTAGAAAGAGTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGGTGGCCCTTTGCTGTAGAAAGAGTTTCTTCTCAGGA 844
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGGCTTATTCAGGACTCATTCTTCC 660
Db 845 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGGCTTATTCAGGACTCATTCTTCC 904
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTCCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTCCAA 964
QY 721 TACTTAGTAATAAGCCTTCAGAAAGAGGCTCAGGAGAGATCATTTGTTGATGCTGTCAA 780
Db 964 TACTTAGTAATAAGCCTTCAGAAAGAGGCTCAGGAGAGATCATTTGTTGATGCTGTCAA 780
```


Db 965 TACTTAGTAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGTCAAA 1024
QY 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCATTCGAATGAATTCGATT 840
Db 1025 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCATTCGAATGAATTCGATT 1084
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Db 1085 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 1144
QY 901 AAGGTTTTTCAGGCGAGAAAATCCTTTTGAATTTATGGAATAAATTTTTCAGAGGAAAA 960
Db 1145 AAGGTTTTTCAGGCGAGAAAATCCTTTTGAATTTATGGAATAAATTTTTCAGAGGAAAA 1204
QY 961 ACAAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGAGTATGCGAGAAACC 1020
Db 1205 ACAAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGAGTATGCGAGAAACC 1264
QY 1021 ACAGATAACGCTTTCACCTTTGGATGAGATTTT 1053
Db 1265 ACAGATAACGCTTTCACCTTTGGATGAGATTTT 1297

RESULT 6

US-11-088-686-44

; Sequence 44, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-088-686-44

Query Match 99.8%; Score 1051.4; DB 27; Length 4955;
Best Local Similarity 99.9%; Pred. No. S.1e-270;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCCGAAGCGCGAAGCGCGCGCTGGATCAGATGAGAGATCATCTTCA 60
Db 245 ATGGGCGACCCCGAAGCGCGAAGCGCGCGCTGGATCAGATGAGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 305 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
QY 121 CGTTTGTTCATCTTCCAAATCAGTACCTGTATTTGGAAAATGTATTAACAGGCACAG 180
Db 365 CGTTTGTTCATCTTCCAAATCAGTACCTGTATTTGGAAAATGTATTAACAGGCACAG 424
QY 181 GCTTCTCTTCTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCAGTGAACAG 240
Db 425 GCTTCTCTTCTGACAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCAGTGAACAG 484
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCTCAATCTTAGCCCTTTTTCAGCCAGTGTAT 300
Db 485 CTTAAAGCAGATGAGAAGTACTTCTCTCAATCTTAGCCCTTTTTCAGCCAGTGTAT 544
QY 301 GGAATTTGAATGAATAATTTGTGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Db 545 GGAATTTGAATGAATAATTTGTGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 604
QY 361 CGCTGTTCTATGGCTTTTCAAAATCTTCATCAGAGATGTTTCACTCAGAGATGATACAGTTTG 420
Db 605 CGCTGTTCTATGGCTTTTCAAAATCTTCATCAGAGATGTTTCACTCAGAGATGATACAGTTTG 664

QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAATTTTATTTAATGCAATTGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAATTTTATTTAATGCAATTGAA 724
QY 481 ACCATGCCCTATGTATTAAGAAAAAGCAGATTGGCCCTTGCATGATGATAGCAGATAGAAA 540
Db 725 ACCATGCCCTATGTATTAAGAAAAAGCAGATTGGCCCTTGCATGATGATAGCAGATAGAAA 784
QY 541 TCTACTTTTGGGGAAGAGTGTGGCCCTTTCCTCTCTGTAGAGAGATTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGGAAGAGTGTGGCCCTTTCCTCTCTGTAGAGAGATTTTCTTCTCAGGA 844
QY 601 TCTTTTGTCTATATTTCTGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Db 845 TCTTTTGTCTATATTTCTGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 904
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTTCACCTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGGACTTTCACCTGTGACTTTGCTTGCCTGATGTTCCAA 964
QY 721 TACTTAGTAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCTCAA 780
Db 965 TACTTAGTAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCTCAA 1024
QY 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCATTCGAATGAATTCGATT 840
Db 1025 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCATTCGAATGAATTCGATT 1084
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATTCTCA 900
Db 1085 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATTCTCA 1144
QY 901 AAGGTTTTTCAGGCGAGAAAATCCTTTTGAATTTATGGAATAAATTTTTCAGAGGAAAA 960
Db 1145 AAGGTTTTTCAGGCGAGAAAATCCTTTTGAATTTATGGAATAAATTTTTCAGAGGAAAA 1204
QY 961 ACAAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGAGTATGCGAGAAACC 1020
Db 1205 ACAAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGAGTATGCGAGAAACC 1264
QY 1021 ACAGATAACGCTTTCACCTTTGGATGAGATTTT 1053
Db 1265 ACAGATAACGCTTTCACCTTTGGATGAGATTTT 1297

RESULT 7

US-10-990-328A-49

; Sequence 49, Application US/10990328A
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328A
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 1849
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-990-328A-49

Query Match 99.8%; Score 1051; DB 15; Length 1849;

Best Local Similarity 99.8%; Pred. No. 4.9e-270;

Matches 1051; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCCGAAGCGCGAAGCGCGCGCTGGATCAGAGTATGAGATCATCTTCA 60
Db 85 ATGGGCGACCCCGAAGCGCGAAGCGCGCGCTGGATCAGAGTATGAGATCATCTTCA 144

```
QY 61 GACACAAAGGAAATGAATGAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 145 GACACAAAGGAAATGAATGAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 204
QY 121 CGGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG 180
Db 205 CGGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG 264
QY 181 GCTTCTCTCTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCTCACTGGAACAAG 240
Db 265 GCTTCTCTCTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCTCACTGGAACAAG 324
QY 241 CTTAAAGCAGATGAGAATCTTCACTCTCAATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db 325 CTTAAAGCAGATGAGAATCTTCACTCTCAATCTTAGCCCTTTTTCAGCCAGTGAT 384
QY 301 GGAATTTGAATGAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Db 385 GGAATTTGAATGAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 444
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 420
Db 445 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 504
QY 421 CTGATAGACACTTACATGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
Db 505 CTGATAGACACTTACATGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 564
QY 481 ACCATGCCCTNTGTTAAGAAAAAGCAGATTTGGCCCTTCGGATGGATAGCAGATAGAAA 540
Db 565 ACCATGCCCTNTGTTAAGAAAAAGCAGATTTGGCCCTTCGGATGGATAGCAGATAGAAA 624
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAAGAGGAGTTTCTTCTCAGGA 600
Db 625 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAAGAGGAGTTTCTTCTCAGGA 684
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAAAGAGAGGTTCTTATGCCAGGACTCACTTTTCC 660
Db 685 TCTTTTGTGCTATATTTCTGGCTTAAAGAAAGAGAGGTTCTTATGCCAGGACTCACTTTTCC 744
QY 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 720
Db 745 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 804
QY 721 TACTTAGTAATAAGCTTTCAGAGAAAGGTCAGGAGATCATTTGATGCTGTCAAA 780
Db 805 TACTTAGTAATAAGCTTTCAGAGAAAGGTCAGGAGATCATTTGATGCTGTCAAA 864
QY 781 ATTGAGCAGGAGTTTAAACAGAAAGCTTTGCCAGTTGGCTCATTTGGAATGCAATT 840
Db 865 ATTGAGCAGGAGTTTAAACAGAAAGCTTTGCCAGTTGGCTCATTTGGAATGCAATT 924
QY 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTTCA 900
Db 925 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTTCA 984
QY 901 AAGGTTTTTCCGCGAGAAATCTTTTGAATTTTATGGAACATTTCTTTAGAGAA 960
Db 985 AAGGTTTTTCCGCGAGAAATCTTTTGAATTTTATGGAACATTTCTTTAGAGAA 1044
QY 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGTATGCGAGAAACC 1020
Db 1045 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGTATGCGAGAAACC 1104
QY 1021 ACAGATAACGCTTCTCACTTGGATGAGATTTT 1053
Db 1105 ACAGATAACGCTTCTCACTTGGATGAGATTTT 1137
```

RESULT 8
PCT-US05-09639-18
; Sequence 18, Application PC/TUS0509639
; GENERAL INFORMATION:

```
; APPLICANT: Yen, Yun  
; TITLE OF INVENTION: DRUG SCREENING  
; FILE REFERENCE: 14037-004W01  
; CURRENT APPLICATION NUMBER: PCT/US05/09639  
; CURRENT FILING DATE: 2005-03-23  
; PRIOR APPLICATION NUMBER: US 60/556,836  
; PRIOR FILING DATE: 2004-03-25  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 18  
; LENGTH: 1056  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
PCT-US05-09639-18  
  
Query Match 99.7%; Score 1049, 8; DB 1; Length 1056;  
Best Local Similarity 99.8%; Pred. No. 8.8e-270; Indels 0; Gaps 0;  
Matches 1051; Conservative 0; Mismatches 2;  
  
QY 1 ATGGGCGACCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60  
Db 1 ATGGGCGACCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60,  
QY 61 GACACAAAGGAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120  
Db 61 GACACAAAGGAGTGAATTAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120  
QY 121 CGGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG 180  
Db 121 CGGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG 180  
QY 181 GCTTCTCTCTGACAGCAGAGAGGTTGATCTATCAAAAGGATCTCCCTCACTGGAACAAG 240  
Db 181 GCTTCTCTCTGACAGCAGAGAGGTTGATCTATCAAAAGGATCTCCCTCACTGGAACAAG 240  
QY 241 CTTAAAGCAGATGAGAATCTTCACTCTCAATCTTAGCCCTTTTTCAGCCAGTGAT 300  
Db 241 CTTAAAGCAGATGAGAATCTTCACTCTCAATCTTAGCCCTTTTTCAGCCAGTGAT 300  
QY 301 GGAATTTGAATGAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360  
Db 301 GGAATTTGAATGAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360  
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420  
Db 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420  
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480  
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480  
QY 481 ACCATGCCCTNTGTTAAGAAAAAGCAGATTTGGCCCTTCGGATGGATAGCAGATAGAAA 540  
Db 481 ACCATGCCCTNTGTTAAGAAAAAGCAGATTTGGCCCTTCGGATGGATAGCAGATAGAAA 540  
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600  
Db 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600  
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAAAGAGAGGTTCTTATGCGAGGACTCACTTTTTC 660  
Db 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAAAGAGAGGTTCTTATGCGAGGACTCACTTTTTC 660  
QY 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTGCTGCTGATGTTTCCAA 720  
Db 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTGCTGCTGATGTTTCCAA 720  
QY 721 TACTTAGTAATAAGCTTTCAGAGAAAGGTCAGGAGATCATTTGATGCTGTCAAA 780  
Db 721 TACTTAGTAATAAGCTTTCAGAGAAAGGTCAGGAGATCATTTGATGCTGTCAAA 780  
QY 781 ATTGAGCAGGAGTTTAAACAGAAAGCTTTGCCAGTTGGCTCATTTGGAATGCAATT 840
```

Db 781 ATTGACGAGGATTTTAAACAGAGCCTTCCAGTGTGCGCTCATTTGAATGAATTCATT 840
Qy 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATCTCA 900
Db 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATCTCA 900
Qy 901 AAGTTTTCAGGACAGAAATCCTTTTGAATTTATGGAACCAATTTCTTTAGAGGAAAA 960
Db 901 AAGTTTTCAGGACAGAAATCCTTTTGAATTTATGGAACCAATTTCTTTAGAGGAAAA 960
Qy 961 ACAAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGGTTTTCAGTTATGCGAGAAACC 1020
Db 961 ACAAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGGTTTTCAGTTATGCGAGAAACC 1020
Qy 1021 ACAGATAACCTCTTCACTTGGATGAGATTTT 1053
Db 1021 ACAGATAACCTCTTCACTTGGATGAGATTTT 1053

RESULT 9
US-11-088-686-18
; Sequence 18, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-088-686-18

Query Match 99.7%; Score 1049.8; DB 27; Length 1056;
Best Local Similarity 99.8%; Pred. No. 8.8e-270;
Matches 1051; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGGCGACCCGGAAGCGCGGAGCGGCGGCTGGATCAGAGATCATCTTCA 60
Db 1 ATGGGCGACCCGGAAGCGCGGAGCGGCGGCTGGATCAGAGATCATCTTCA 60
Qy 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGCCACTCTTAAGAAAGAGTTCTCG 120
Db 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGCCACTCTTAAGAAAGAGTTCTCG 120
Qy 121 CGTTTGTCACTTTCATTCAGTACCTGTATTTGGAATGTAATAACAGGCACAG 180
Db 121 CGTTTGTCACTTTCATTCAGTACCTGTATTTGGAATGTAATAACAGGCACAG 180
Qy 181 GCTTCCTCTGACACAGAGGTTGACTTATCAAGGATCTCCCTCACTGACACAG 240
Db 181 GCTTCCTCTGACACAGAGGTTGACTTATCAAGGATCTCCCTCACTGACACAG 240
Qy 241 CTTAAGCAGATGAGAAGTACTTCACTCTCAGATCTAGGCTTTTTCAGGCGAGTAT 300
Db 241 CTTAAGCAGATGAGAAGTACTTCACTCTCAGATCTAGGCTTTTTCAGGCGAGTAT 300
Qy 301 GGAATTTGTAATGAAATTTGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGCT 360
Db 301 GGAATTTGTAATGAAATTTGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGCT 360
Qy 361 CGCTGTTCTATGCGCTTCAATTTCTATCGAGAATTTTCACTCAGAGATGATACGTTG 420
Db 361 CGCTGTTCTATGCGCTTCAATTTCTATCGAGAATTTTCACTCAGAGATGATACGTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGAA 480

Qy 481 ACCATGCCCTATGTTAAGAAAAACGAGTTGGGCTTCCGATGATAGCAGATGAAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAACGAGTTGGGCTTCCGATGATAGCAGATGAAAA 540
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTCTGCTGTAGAGAGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTCTGCTGTAGAGAGAGTTTCTTCTCAGGA 600
Qy 601 TCTTTTCTGCTATATTTCTGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
Db 601 TCTTTTCTGCTATATTTCTGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
Qy 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGATGTTCCAA 720
Qy 721 TACTTAGTAAATAAGCCTTTCAGAAAGAGGCTCAGGAGATCATTTGTGATGCTGTCAA 780
Db 721 TACTTAGTAAATAAGCCTTTCAGAAAGAGGCTCAGGAGATCATTTGTGATGCTGTCAA 780
Qy 781 ATTGACGAGGATTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATTCATT 840
Db 781 ATTGACGAGGATTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATTCATT 840
Qy 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGTGGAACCTTGGATCTCA 900
Db 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGTGGAACCTTGGATCTCA 900
Qy 901 AAGTTTTCAGGCGAGAAAAATCCTTTTGAATTTATGGAACCAATTTCTTTAGAGGAAAA 960
Db 901 AAGTTTTCAGGCGAGAAAAATCCTTTTGAATTTATGGAACCAATTTCTTTAGAGGAAAA 960
Qy 961 ACAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1020
Db 961 ACAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1020
Qy 1021 ACAGATAACCTCTTCACTTGGATGAGATTTT 1053
Db 1021 ACAGATAACCTCTTCACTTGGATGAGATTTT 1053

RESULT 10
PCT-US05-09639-20
; Sequence 20, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-20

Query Match 99.5%; Score 1048.2; DB 1; Length 1056;
Best Local Similarity 99.7%; Pred. No. 2.4e-269;
Matches 1050; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGGGCGACCCGGAAGCGCGGAGCGGCTGGATCAGAGATCAGATCATCTTCA 60
Db 1 ATGGGCGACCCGGAAGCGCGGAGCGGCTGGATCAGAGATCAGATCATCTTCA 60
Qy 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGCCACTCTTAAGAAAGAGTTCTCG 120
Db 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGCCACTCTTAAGAAAGAGTTCTCG 120

```
QY 121 CGTTTGTCAATCTTCCAAATCCAGTACCCTGATATTTGGAAATGATATAACAGGCACAG 180
Db |||||||
QY 121 CGTTTGTCAATCTTCCAAATCCAGTACCCTGATATTTGGAAATGATATAACAGGCACAG 180
Db |||||||
QY 181 GCTTCTTCTGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGAAACAG 240
Db |||||||
QY 181 GCTTCTTCTGACAGCAGAGAGGTGCGACTTATCAAGGATCTCCCTCACTGGAAACAG 240
Db |||||||
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db |||||||
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db |||||||
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTCCAGGTTCCAGAGGCT 360
Db |||||||
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTCCAGGTTCCAGAGGCT 360
Db |||||||
QY 361 CGCTGTTCTATGGCTTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 420
Db |||||||
QY 361 CGCTGTTCTATGGGCTTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 420
Db |||||||
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTTGAA 480
Db |||||||
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTTGAA 480
Db |||||||
QY 481 ACCATGCCCTATGTTAAAGAAAGCAGATTTGGGCCCTTCGATGGATAGCAGATAGAAA 540
Db |||||||
QY 481 ACCATGCCCTATGTTAAAGAAAGCAGATTTGGGCCCTTCGATGGATAGCAGATAGAAA 540
Db |||||||
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGATTTTCTTCTCAGGA 600
Db |||||||
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGATTTTCTTCTCAGGA 600
Db |||||||
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCCAGACTCACTTTTTC 660
Db |||||||
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCCAGACTCACTTTTTC 660
Db |||||||
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGTGCTGATGTTCCAA 720
Db |||||||
QY 721 TACTTAGTAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db |||||||
QY 721 TACTTAGTAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db |||||||
QY 781 ATTGAGCAGGAGTTTTAAACAGAGCCTTGGCAGTTGGCCTCATTTGGAATGCAATT 840
Db |||||||
QY 781 ATTGAGCAGGAGTTTTAAACAGAGCCTTGGCAGTTGGCCTCATTTGGAATGCAATT 840
Db |||||||
QY 841 TTGATGAAACAGTACATTTAGTGTGATGCTGACAGATTTAGTGGAACTTGGATTCTCA 900
Db |||||||
QY 841 TTGATGAAACAGTACATTTAGTGTGATGCTGACAGATTTAGTGGAACTTGGATTCTCA 900
Db |||||||
QY 901 AAGGTTTTTCAGGCAGAAAATCTTTGATTTATGGAATAACATTTCTTTAGAGGAAA 960
Db |||||||
QY 901 AAGGTTTTTCAGGCAGAAAATCTTTGATTTATGGAATAACATTTCTTTAGAGGAAA 960
Db |||||||
QY 961 ACAAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTGCAGTTATGCCAGAAAC 1020
Db |||||||
QY 961 ACAAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTGCAGTTATGCCAGAAAC 1020
Db |||||||
QY 1021 ACAGATAACGCTTCCACCTTGGATGAGATTTT 1053
Db |||||||
QY 1021 ACAGATAACGCTTCCACCTTGGATGAGATTTT 1053
Db |||||||
```

RESULT 11

```
US-11-088-686-20
; Sequence 20, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088, 686
```

```
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-088-686-20
```

```
Query Match 99.5%; Score 1048.2; DB 27; Length 1056;
Best Local Similarity 99.7%; Pred. No. 2.4e-269;
Matches 1050; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 ATCGGCGACCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db |||||||
QY 1 ATCGGCGACCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db |||||||
QY 61 GACACCAACGAAAGTGAATATAAGTCAAATGAAGGCCACTCTCTAAGAAAGATTCCTGC 120
Db |||||||
QY 61 GACACCAACGAAAGTGAATATAAGTCAAATGAAGGCCACTCTCTAAGAAAGATTCCTGC 120
Db |||||||
QY 121 CGGTTTGTCACTTTTCCAAATCCAGTACCTGATATTTGGAAATGATATAACAGGCACAG 180
Db |||||||
QY 121 CGGTTTGTCACTTTTCCAAATCCAGTACCTGATATTTGGAAATGATATAACAGGCACAG 180
Db |||||||
QY 181 GCTTCTTCTGGACAGCAGAGAGGTGACTTATCAAAAGGATCTCCCTCACTGGAAACAG 240
Db |||||||
QY 181 GCTTCTTCTGGACAGCAGAGAGGTGACTTATCAAAAGGATCTCCCTCACTGGAAACAG 240
Db |||||||
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db |||||||
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db |||||||
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTCCAGAGGCT 360
Db |||||||
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTCCAGAGGCT 360
Db |||||||
QY 361 CGCTGTTCTATGGCTTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 420
Db |||||||
QY 361 CGCTGTTCTATGGGCTTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 420
Db |||||||
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTTGAA 480
Db |||||||
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTTGAA 480
Db |||||||
QY 481 ACCATGCCCTATGTTAAAGAAAGCAGATTTGGGCCCTTCGATGGATAGCAGATAGAAA 540
Db |||||||
QY 481 ACCATGCCCTATGTTAAAGAAAGCAGATTTGGGCCCTTCGATGGATAGCAGATAGAAA 540
Db |||||||
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGATTTTCTTCTCAGGA 600
Db |||||||
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGATTTTCTTCTCAGGA 600
Db |||||||
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCCAGACTCACTTTTTC 660
Db |||||||
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCCAGACTCACTTTTTC 660
Db |||||||
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGTGCTGATGTTCCAA 720
Db |||||||
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGTGCTGATGTTCCAA 720
Db |||||||
QY 721 TACTTAGTAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db |||||||
QY 721 TACTTAGTAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db |||||||
QY 781 ATTGAGCAGGAGTTTTAAACAGAGCCTTGGCAGTTGGCCTCATTTGGAATGCAATT 840
Db |||||||
QY 781 ATTGAGCAGGAGTTTTAAACAGAGCCTTGGCAGTTGGCCTCATTTGGAATGCAATT 840
Db |||||||
QY 841 TTGATGAAACAGTACATTTAGTGTGATGCTGACAGATTTAGTGGAACTTGGATTCTCA 900
Db |||||||
```

```
Db 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 900
Qy 901 AAGTTTTTTCAGCAGAAAATCCTTTTGTATTTATGAAAAATTTCTTTAGAAAGAAA 960
Db 901 AAGTTTTTTCAGCAGAAAATCCTTTTGTATTTATGAAAAATTTCTTTAGAAAGAAA 960
Qy 961 ACAAATTTCTTTGAGAAAAGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
Db 961 ACAAATTTCTTTGAGAAAAGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
Qy 1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053

RESULT 12
PCT-US05-09639-22
; Sequence 22, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-22

Query Match 99.4%; Score 1046.6; DB 1; Length 1056;
Best Local Similarity 99.6%; Pred. No. 6.3e-269;
Matches 1049; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGGGGGACCCGGAAGGCGGAGGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGGGGACCCGGAAGGCGGAGGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Qy 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Qy 121 CGTTTGTCTCATCTTTCCAAATCCAGTACCCTGTATTTGGAAAAATGTATAAACAGGCACAG 180
Db 121 CGTTTGTCTCATCTTTCCAAATCCAGTACCCTGTATTTGGAAAAATGTATAAACAGGCACAG 180
Qy 181 GCTTCCTCTTGGACAGAGAGGTTGACTTATCAAGATCTCCCTCACTGGAACAAG 240
Db 181 GCTTCCTCTTGGACAGAGAGGTTGACTTATCAAGATCTCCCTCACTGGAACAAG 240
Qy 241 CTTAAAGCAGATGAGAGTACTTCTCATCTTCACTTACCTTTTTCGAGCCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAGTACTTCTCATCTTCACTTACCTTTTTCGAGCCAGTGAT 300
Qy 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Qy 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAATTTTCTCATCAGAGTGTACAGTTTG 420
Db 361 CGCTGTTTCTTGGCTTTCAAAATTTCTCATCGAATTTTCTCATCAGAGTGTATTTAGTTTG 420
Qy 421 CTGATAGACATTAATCAGAGATCCCAAGAAAAGGGAATTTTATTTATGCAATTGAA 480
Db 421 CTGATAGACATTAATCAGAGATCCCAAGAAAAGGGAATTTTATTTATGCAATTGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAAGAGAGATTTGGGCTTTCGATGGATAGCAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAGAGAGATTTGGGCTTTCGATGGATAGCAGATAGAAA 540
```

```
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCCTTTGCTGCTGTAGAAAGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCCTTTGCTGCTGTAGAAAGAGTTTCTTCTCAGGA 600
Qy 601 TCTTTTGTCTCTATATTTCTGGCTAAAGAGAGGTCTTATGCCAGGACTCATCTTTTCC 660
Db 601 TCTTTTGTCTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCATCTTTTCC 660
Qy 661 AATGAACTCATCAGCAGAGATGAAGACTTTCACCTGTGACCTTTGCTGTGCTGATGTTCCAA 720
Db 661 AATGAACTCATCAGCAGAGATGAAGACTTTCACCTGTGACCTTTGCTGTGCTGATGTTCCAA 720
Qy 721 TACTTAGTAATAAGCCTTCAGAAAGAGGTGAGGAGATCATTTGTGATGCTGTCAAA 780
Db 721 TACTTAGTAATAAGCCTTCAGAAAGAGGTGAGGAGATCATTTGTGATGCTGTCAAA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCTTCCAGTTGGCCTCATTTGGAATGAATTCAT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGCCTTCCAGTTGGCCTCATTTGGAATGAATTCAT 840
Qy 841 TTGATGAACAGTACATTTAGTTTGTAGCTGACAGATTTACTTTGGAACTTTGGATTCTCA 900
Db 841 TTGATGAACAGTACATTTAGTTTGTAGCTGACAGATTTACTTTGGAACTTTGGATTCTCA 900
Qy 901 AAGTTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGAAAACATTTCTTTAGAAAGAAA 960
Db 901 AAGTTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGAAAACATTTCTTTAGAAAGAAA 960
Qy 961 ACAAATTTCTTTGAGAAAAGAGTTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
Db 961 ACAAATTTCTTTGAGAAAAGAGTTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
Qy 1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053

RESULT 13
PCT-US05-09639-24
; Sequence 24, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-24

Query Match 99.4%; Score 1046.6; DB 1; Length 1056;
Best Local Similarity 99.6%; Pred. No. 6.3e-269;
Matches 1049; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGGGGGACCCGGAAGGCGGAGGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGGGGACCCGGAAGGCGGAGGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Qy 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Qy 121 CGTTTGTCTCATCTTTCCAAATCCAGTACCCTGTATTTGGAAAAATGTATAAACAGGCACAG 180
Db 121 CGTTTGTCTCATCTTTCCAAATCCAGTACCCTGTATTTGGAAAAATGTATAAACAGGCACAG 180
Qy 181 GCTTCCTCTTGGACAGAGAGGTTGACTTATCAAGATCTCCCTCACTGGAACAAG 240
Db 181 GCTTCCTCTTGGACAGAGAGGTTGACTTATCAAGATCTCCCTCACTGGAACAAG 240
Qy 241 CTTAAAGCAGATGAGAGTACTTCTCATCTTCACTTACCTTTTTCGAGCCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAGTACTTCTCATCTTCACTTACCTTTTTCGAGCCAGTGAT 300
Qy 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Qy 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAATTTTCTCATCAGAGTGTACAGTTTG 420
Db 361 CGCTGTTTCTTGGCTTTCAAAATTTCTCATCGAATTTTCTCATCAGAGTGTATTTAGTTTG 420
Qy 421 CTGATAGACATTAATCAGAGATCCCAAGAAAAGGGAATTTTATTTATGCAATTGAA 480
Db 421 CTGATAGACATTAATCAGAGATCCCAAGAAAAGGGAATTTTATTTATGCAATTGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAAGAGAGATTTGGGCTTTCGATGGATAGCAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAGAGAGATTTGGGCTTTCGATGGATAGCAGATAGAAA 540
```

```
181 GCTTCCTCTGACAGCAGAGAGGTGACTTATCAAGAGATCCCTCACTGGAACAAG 240
Db      |||
181 GCTTCCTCTGACAGCAGAGAGGTGACTTATCAAGAGATCCCTCACTGGAACAAG 240
Qy      |||
241 CTTAAAGCAGATGAGAGTACTTCTCTCATCTTACATCTTAGCCCTTTTTCGACGCAATGAT 300
Db      |||
241 CTTAAAGCAGATGAGAGTACTTCTCTCATCTTACATCTTAGCCCTTTTTCGACGCAATGAT 300
Qy      |||
301 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
Db      |||
301 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
Qy      |||
361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db      |||
361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGAATGTTCACTCAGAGATGTGAGTTTG 420
Qy      |||
421 CTGATAGACACTTATCATGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTGAA 480
Db      |||
421 CTGATAGACACTTATCATGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTGAA 480
Qy      |||
481 ACCATGCCCTATGTTAAGAAATAAGCAGATTTGGGCGCTTGCATGGATAGCAGATAGAAA 540
Db      |||
481 ACCATGCCCTATGTTAAGAAATAAGCAGATTTGGGCGCTTGCATGGATAGCAGATAGAAA 540
Qy      |||
541 TCTACTTTTGGGAAAGAGTGGTGGCGCTTTGCTGTAGAGAGGTTTCTTCTCAGGA 600
Db      |||
541 TCTACTTTTGGGAAAGAGTGGTGGCGCTTTGCTGTAGAGAGGTTTCTTCTCAGGA 600
Qy      |||
601 TCTTTTGTCTATATTTCTGGCTTAAAGAAAGAGAGTCTTATGCGAGGACTCACTTTTTC 660
Db      |||
601 TCTTTTGTCTATATTTCTGGCTTAAAGAAAGAGAGTCTTATGCGAGGACTCACTTTTTC 660
Qy      |||
661 AATGAACCTCATGAGAGATGAGGACTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db      |||
661 AATGAACCTCATGAGAGATGAGGACTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Qy      |||
721 TACTTAGTAAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAA 780
Db      |||
721 TACTTAGTAAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAA 780
Qy      |||
781 ATTGACGAGGTTTAAACAGAGCCTTGCCAGTTGGCTCATTTGGAATGAATTGCATT 840
Db      |||
781 ATTGACGAGGTTTAAACAGAGCCTTGCCAGTTGGCTCATTTGGAATGAATTGCATT 840
Qy      |||
841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATTCTCA 900
Db      |||
841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATTCTCA 900
Qy      |||
901 AAGGTTTTTCAGGCAAGAAATCCTTTTGAATTTTATGGAACCAATTTCTTTAGAGGAAA 960
Db      |||
901 AAGGTTTTTCAGGCAAGAAATCCTTTTGAATTTTATGGAACCAATTTCTTTAGAGGAAA 960
Qy      |||
961 ACAAAATTTCTTTAGAAACAGGTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1020
Db      |||
961 ACAAAATTTCTTTAGAAACAGGTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1020
Qy      |||
1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
Db      |||
1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
```

RESULT 14

```
US-11-088-686-22
; Sequence 22, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-11-088-686-22
```

Query Match

```
99.4%; Score 1046.6; DB 27; Length 1056;
Best Local Similarity 99.6%; Pred. No. 6.3e-269;
Matches 1049; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
Qy      1 ATGGCGACCCGGAAGCGCGGCTGATCAGGATGAGAGATCACTTTCA 60
Db      |||
Qy      1 ATGGCGACCCGGAAGCGCGGCTGATCAGGATGAGAGATCACTTTCA 60
Db      |||
Qy      61 GACACCAAGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGTTCTCGC 120
Db      |||
Qy      61 GACACCAAGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGTTCTCGC 120
Db      |||
Qy      121 CGGTTTGTCTATCTTTCCAAATCCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG 180
Db      |||
Qy      121 CGGTTTGTCTATCTTTCCAAATCCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG 180
Db      |||
Qy      181 GCTTCTCTTCTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db      |||
Qy      181 GCTTCTCTTCTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db      |||
Qy      241 CTTAAAGCAGATGAGAGTACTTCTCATCTCAGATCTTAGCCTTTTTCGACGCAATGAT 300
Db      |||
Qy      241 CTTAAAGCAGATGAGAGTACTTCTCATCTCAGATCTTAGCCTTTTTCGACGCAATGAT 300
Db      |||
Qy      301 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
Db      |||
Qy      301 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
Db      |||
Qy      361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGAATGTTCTCAGAGATGTACAGTTTG 420
Db      |||
Qy      361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGAATGTTCTCAGAGATGTAGTTTG 420
Db      |||
Qy      421 CTGATAGACACTTATCATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTGAA 480
Db      |||
Qy      421 CTGATAGACACTTATCATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTGAA 480
Db      |||
Qy      481 ACCATGCCCTATGTTAAGAAATAAGCAGATTTGGGCGCTTGCATGGATAGCAGATAGAAA 540
Db      |||
Qy      481 ACCATGCCCTATGTTAAGAAATAAGCAGATTTGGGCGCTTGCATGGATAGCAGATAGAAA 540
Db      |||
Qy      541 TCTACTTTTGGGAAAGAGTGGTGGCGCTTTGCTGTAGAGAGGTTTCTTCTCAGGA 600
Db      |||
Qy      541 TCTACTTTTGGGAAAGAGTGGTGGCGCTTTGCTGTAGAGAGGTTTCTTCTCAGGA 600
Db      |||
Qy      601 TCTTTTGTCTATATTTCTGGCTTAAAGAAAGAGGCTTATGCGAGGACTCACTTTTTC 660
Db      |||
Qy      601 TCTTTTGTCTATATTTCTGGCTTAAAGAAAGAGGCTTATGCGAGGACTCACTTTTTC 660
Db      |||
Qy      661 AATGAACCTCATGAGAGATGAGGACTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db      |||
Qy      661 AATGAACCTCATGAGAGATGAGGACTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db      |||
Qy      721 TACTTAGTAAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAA 780
Db      |||
Qy      721 TACTTAGTAAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAA 780
Db      |||
Qy      781 ATTGACGAGGTTTAAACAGAGCCTTGCCAGTTGGCTCATTTGGAATGAATTGCATT 840
Db      |||
Qy      781 ATTGACGAGGTTTAAACAGAGCCTTGCCAGTTGGCTCATTTGGAATGAATTGCATT 840
Db      |||
Qy      841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATTCTCA 900
Db      |||
Qy      841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATTCTCA 900
Db      |||
Qy      901 AAGGTTTTTCAGGCAAGAAATCCTTTTGAATTTTATGGAACCAATTTCTTTAGAGGAAA 960
Db      |||
Qy      |||
```


Db 901 AAGTTTTTCAGCAGAGAAATCCTTTTGATTTTATGGAAGAAACATTTCTTTAGAGAGAAA 960
Qy 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
Qy 1021 ACAGATAAGCTCTTACCTTTGGATGAGATTTT 1053
Db 1021 ACAGATAAGCTCTTACCTTTGGATGAGATTTT 1053

RESULT 15
US-11-088-686-24
; Sequence 24, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yeh, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-11-088-686-24

Query Match 99.4%; Score 1046.6; DB 27; Length 1056;
Best Local Similarity 99.6%; Pred. No. 6.3e-269;
Matches 1049; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGGGCGACCCGGAAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
Db 1 ATGGGCGACCCGGAAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
Qy 61 GACACCAAGCAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAAGCAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Qy 121 CGTTTGTCTATCTTTCCATCCAGTACCTGATATTTTGGAAATGATTAACAGGACACAG 180
Db 121 CGTTTGTCTATCTTTCCATCCAGTACCTGATATTTTGGAAATGATTAACAGGACACAG 180
Qy 181 GCTTCCTCTTGGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
Db 181 GCTTCCTCTTGGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
Qy 241 CTTAAAGCAGATGAGAGTACTTCTCATCATCTTACCTTTTGGAGCCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAGTACTTCTCATCATCTTACCTTTTGGAGCCAGTGAT 300
Qy 301 GGAATTTGAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCAGAGGCT 360
Db 301 GGAATTTGAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCAGAGGCT 360
Qy 361 CGCTGTTTCTATGGCTTTTCAATCTCATCGAGATGTTTCACTCAGAGATGTACAGTTG 420
Db 361 CGCTGTTTCTATGGCTTTTCAATCTCATCGAGATGTTTCACTCAGAGATGTACAGTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCCCTTCGGATGGATAGCAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCCCTTCGGATGGATAGCAGATAGAAA 540
Qy 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTTGTCTGTAGAGAGGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTTGTCTGTAGAGAGGAGTTTCTTCTCAGGA 600

Search completed: October 30, 2005, 06:01:28
Job time : 1375.5 secs

Qy 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACATTTTTC 660
Db 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACATTTTTC 660
Qy 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGCTGATGTTCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGCTGATGTTCAA 720
Qy 721 TACTTAGTAATTAAGCCCTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db 721 TACTTAGTAATTAAGCCCTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAGAGCTTGGCAGTTGGCCCTCATTTGGAATGAATTCATT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGAGCTTGGCAGTTGGCCCTCATTTGGAATGAATTCATT 840
Qy 841 TTGATGAACACAGTACATTGAGTTTGTAGCTGACAGATTTCTTTGGAACTTTGGATTTCTCA 900
Db 841 TTGATGAACACAGTACATTGAGTTTGTAGCTGACAGATTTCTTTGGAACTTTGGATTTCTCA 900
Qy 901 AAGTTTTTTCAGGAGAGAAATCCTTTTGATTTTATGGAACATTTCTTTAGAGAGAAA 960
Db 901 AAGTTTTTTCAGGAGAGAAATCCTTTTGATTTTATGGAACATTTCTTTAGAGAGAAA 960
Qy 961 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTATGSCAGAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTATGSCAGAAACC 1020
Qy 1021 ACAGATAACGCTCTTTCACCTTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGCTCTTTCACCTTTGGATGAGATTTT 1053

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2005, 06:55:26 ; Search time 177 Seconds
(without alignments)
2316.216 Million cell updates/sec

Title: US-10-698-228-1

Perfect score: 1821
Sequence: 1 MGDPRPEAAGLDQDERSSS.....QRFVMAETTDNVFTLDADF 351

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 6959266 seqs, 1168006243 residues

Total number of hits satisfying chosen parameters: 6959266

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_AA_Main:*

- 1: /cgn2_6/ptodata/1/paa/US06 COMB.pcp.*
- 2: /cgn2_6/ptodata/1/paa/US06 COMB.pcp.*
- 3: /cgn2_6/ptodata/1/paa/US07 COMB.pcp.*
- 4: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 5: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 6: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 7: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 8: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 9: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 10: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 11: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 12: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 13: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 14: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 15: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 16: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 17: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 18: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 19: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 20: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 21: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 22: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 23: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 24: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 25: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 26: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 27: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 28: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 29: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 30: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 31: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 32: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 33: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 34: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 35: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 36: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 37: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1821	100.0	351	22	US-09-787-491B-11	Sequence 11, Appl
2	1821	100.0	351	22	US-09-791-537-77474	Sequence 77474, A
3	1821	100.0	351	27	US-10-170-205E-19045	Sequence 19045, A
4	1821	100.0	351	32	US-10-698-228-1	Sequence 1, Appl
5	1821	100.0	351	35	US-10-990-328-7547	Sequence 7547, Ap
6	1821	100.0	351	37	US-60-128-660-7	Sequence 7, Appl
7	1821	100.0	351	37	US-60-449-629-94	Sequence 94, Appl
8	1821	100.0	351	37	US-60-505-218-322	Sequence 322, App
9	1821	100.0	366	22	US-09-791-537-27671	Sequence 27671, A
10	1817	99.8	351	22	US-09-629-469A-12811	Sequence 12811, A
11	1817	99.8	351	22	US-09-791-537-112185	Sequence 112185, A
12	1817	99.8	351	35	US-10-917-503-12811	Sequence 12811, A
13	1454	79.8	389	1	PCT-US02-10824-143	Sequence 143, App
14	1454	79.8	389	1	PCT-US03-17409-171	Sequence 171, App
15	1454	79.8	389	10	US-08-663-617-33	Sequence 33, Appl
16	1454	79.8	389	12	US-08-837-599-33	Sequence 33, Appl
17	1454	79.8	389	22	US-09-791-537-88179	Sequence 88179, A
18	1454	79.8	389	27	US-10-170-205E-8018	Sequence 8018, Ap
19	1454	79.8	389	30	US-10-408-765-559	Sequence 559, App
20	1454	79.8	389	30	US-10-408-765A-559	Sequence 559, App
21	1454	79.8	389	32	US-10-698-228-5	Sequence 5, Appl
22	1454	79.8	389	33	US-10-733-878-457	Sequence 457, App
23	1454	79.8	389	33	US-10-756-149-4899	Sequence 4899, Ap
24	1454	79.8	389	35	US-10-940-774-6016	Sequence 6016, Ap
25	1454	79.8	389	37	US-60-389-987-559	Sequence 559, App
26	1454	79.8	389	37	US-60-412-418-559	Sequence 559, App
27	1454	79.8	389	37	US-60-452-680-20616	Sequence 20616, A
28	1454	79.8	389	37	US-60-568-073-984	Sequence 984, App
29	1454	79.8	413	1	PCT-US00-05882-1347	Sequence 1347, Ap
30	1454	79.8	413	24	US-09-925-301-1347	Sequence 1347, Ap
31	1454	79.8	449	35	US-10-990-328-12349	Sequence 12349, A
32	1454	79.8	449	35	US-10-990-328-12350	Sequence 12350, A
33	1454	79.8	453	1	PCT-US01-04926A-686	Sequence 686, App
34	1454	79.8	453	1	PCT-US01-08631-45646	Sequence 45646, A
35	1454	79.8	453	28	US-10-420-335-686	Sequence 686, App
36	1454	79.8	453	28	US-10-290-752-686	Sequence 686, App
37	1454	79.8	453	35	US-10-940-774-7896	Sequence 7896, Ap
38	1443	79.2	390	10	US-08-663-617-35	Sequence 35, Appl
39	1443	79.2	390	12	US-08-837-599-35	Sequence 35, Appl
40	1443	79.2	390	22	US-09-791-537-11257	Sequence 11257, A
41	1425.5	78.3	386	22	US-09-791-537-114288	Sequence 114288, A
42	1425.5	78.3	386	30	US-10-403-571-76	Sequence 76, Appl
43	1425.5	78.3	386	30	US-10-405-887-76	Sequence 76, Appl
44	1421	78.0	386	22	US-09-791-537-36076	Sequence 36076, A
45	1349.5	74.1	399	22	US-09-791-537-1129	Sequence 1129, Ap

ALIGNMENTS

RESULT 1
US-09-787-491B-11
; Sequence 11, Application US/09787491B
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; TANG, Y. Tom;
; APPLICANT: CORLEY, Neil C.; GUEGLER, Karl J.;
; APPLICANT: GORGONE, Gina A.; PATTERSON, Chandra;
; APPLICANT: HILLMAN, Jennifer L.; BAUGHN, Mariah R.;
; APPLICANT: LAL, Preeti; AZIMZAI, Yalda;
; APPLICANT: YUE, Henry; YANG, Junming
; TITLE OF INVENTION: RNA-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0600 USN
; CURRENT APPLICATION NUMBER: US/09/787,491B
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: PCT/US99/21688
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US 60/128,660
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: US 60/069,391

;
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: US 60/183,025
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: US 60/155,246
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: US 09/158,720
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PERL Program
; SEQ ID NO 11
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 2604449CD1
US-09-787-491B-11

Query Match 100.0%; Score 1821; DB 22; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.3e-180;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPIWKMYKQAA 60
DB 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPIWKMYKQAA 60
QY 61 ASFTAAEEVDLSKDLPHWNKLKADEKYFISHILAFPAASDGI VNNELVERFSQEVQVPEA 120
DB 61 ASFTAAEEVDLSKDLPHWNKLKADEKYFISHILAFPAASDGI VNNELVERFSQEVQVPEA 120
QY 121 RCFYGFQI LIENVHSEMYSLIDTYIRDPKKEFLFNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFYGFQI LIENVHSEMYSLIDTYIRDPKKEFLFNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDFACLMFQ 240
DB 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDFACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVGLGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVGLGFS 300
QY 301 KVFOAENPPDFMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 301 KVFOAENPPDFMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 2

US-09-791-537-77474
; Sequence 77474, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 77474
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-77474

Query Match 100.0%; Score 1821; DB 22; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.3e-180;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPIWKMYKQAA 60
DB 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPIWKMYKQAA 60

DB 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPIWKMYKQAA 60
QY 61 ASFTAAEEVDLSKDLPHWNKLKADEKYFISHILAFPAASDGI VNNELVERFSQEVQVPEA 120
DB 61 ASFTAAEEVDLSKDLPHWNKLKADEKYFISHILAFPAASDGI VNNELVERFSQEVQVPEA 120
QY 121 RCFYGFQI LIENVHSEMYSLIDTYIRDPKKEFLFNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFYGFQI LIENVHSEMYSLIDTYIRDPKKEFLFNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDFACLMFQ 240
DB 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDFACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVGLGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVGLGFS 300
QY 301 KVFOAENPPDFMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 301 KVFOAENPPDFMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 3

US-10-170-205E-19045
; Sequence 19045, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19045
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-19045

Query Match 100.0%; Score 1821; DB 27; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.3e-180;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPIWKMYKQAA 60
DB 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPIWKMYKQAA 60
QY 61 ASFTAAEEVDLSKDLPHWNKLKADEKYFISHILAFPAASDGI VNNELVERFSQEVQVPEA 120
DB 61 ASFTAAEEVDLSKDLPHWNKLKADEKYFISHILAFPAASDGI VNNELVERFSQEVQVPEA 120
QY 121 RCFYGFQI LIENVHSEMYSLIDTYIRDPKKEFLFNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFYGFQI LIENVHSEMYSLIDTYIRDPKKEFLFNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDFACLMFQ 240
DB 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDFACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVGLGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVGLGFS 300
QY 301 KVFOAENPPDFMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 301 KVFOAENPPDFMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 4

US-10-698-228-1
; Sequence 1, Application US/10698228

```
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 1
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-1

Query Match      100.0%; Score 1821; DB 32; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.3e-180;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSDDTNESEIKSNEEPRLKSSRRFVFPPIQYDPDIWKYKQAO 60
DB 1 MGDPERPEAAGLDQDERSDDTNESEIKSNEEPRLKSSRRFVFPPIQYDPDIWKYKQAO 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADKEYFISHILAFFAASDGIWVNLVERFSQEVQVPEA 120
DB 61 ASFWTAEVDLSKDLPHWNKLKADKEYFISHILAFFAASDGIWVNLVERFSQEVQVPEA 120
QY 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKREFFLNAIETMPYVKKADWALRWIADRK 180
DB 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKREFFLNAIETMPYVKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGKGLMPLGTFSNELISRDEGLHCDFACLMFQ 240
DB 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGKGLMPLGTFSNELISRDEGLHCDFACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLVVELGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLVVELGFS 300
QY 301 KVFOAENPPDFMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 301 KVFOAENPPDFMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 6
US-60-128-660-7
; Sequence 7, Application US/60128660
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Tang, Y. Tom
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Yang, Junming
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: RNA-ASSOCIATED PROTEINS
; FILE REFERENCES: PF-0682 P
; CURRENT APPLICATION NUMBER: US/60/128,660
; CURRENT FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 2604449
US-60-128-660-7

Query Match      100.0%; Score 1821; DB 37; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.3e-180;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSDDTNESEIKSNEEPRLKSSRRFVFPPIQYDPDIWKYKQAO 60
DB 1 MGDPERPEAAGLDQDERSDDTNESEIKSNEEPRLKSSRRFVFPPIQYDPDIWKYKQAO 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADKEYFISHILAFFAASDGIWVNLVERFSQEVQVPEA 120
DB 61 ASFWTAEVDLSKDLPHWNKLKADKEYFISHILAFFAASDGIWVNLVERFSQEVQVPEA 120
QY 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKREFFLNAIETMPYVKKADWALRWIADRK 180
DB 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKREFFLNAIETMPYVKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGKGLMPLGTFSNELISRDEGLHCDFACLMFQ 240
DB 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGKGLMPLGTFSNELISRDEGLHCDFACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLVVELGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLVVELGFS 300

Query Match      100.0%; Score 1821; DB 35; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.3e-180;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSDDTNESEIKSNEEPRLKSSRRFVFPPIQYDPDIWKYKQAO 60
DB 1 MGDPERPEAAGLDQDERSDDTNESEIKSNEEPRLKSSRRFVFPPIQYDPDIWKYKQAO 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADKEYFISHILAFFAASDGIWVNLVERFSQEVQVPEA 120
DB 61 ASFWTAEVDLSKDLPHWNKLKADKEYFISHILAFFAASDGIWVNLVERFSQEVQVPEA 120
QY 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKREFFLNAIETMPYVKKADWALRWIADRK 180
DB 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKREFFLNAIETMPYVKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGKGLMPLGTFSNELISRDEGLHCDFACLMFQ 240
DB 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGKGLMPLGTFSNELISRDEGLHCDFACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLVVELGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLVVELGFS 300
```

```
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 1
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-1

Query Match      100.0%; Score 1821; DB 32; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.3e-180;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSDDTNESEIKSNEEPRLKSSRRFVFPPIQYDPDIWKYKQAO 60
DB 1 MGDPERPEAAGLDQDERSDDTNESEIKSNEEPRLKSSRRFVFPPIQYDPDIWKYKQAO 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADKEYFISHILAFFAASDGIWVNLVERFSQEVQVPEA 120
DB 61 ASFWTAEVDLSKDLPHWNKLKADKEYFISHILAFFAASDGIWVNLVERFSQEVQVPEA 120
QY 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKREFFLNAIETMPYVKKADWALRWIADRK 180
DB 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKREFFLNAIETMPYVKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGKGLMPLGTFSNELISRDEGLHCDFACLMFQ 240
DB 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGKGLMPLGTFSNELISRDEGLHCDFACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLVVELGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLVVELGFS 300
QY 301 KVFOAENPPDFMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 301 KVFOAENPPDFMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 5
US-10-990-328-7547
; Sequence 7547, Application US/10990328
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7547
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION:
US-10-990-328-7547

Query Match      100.0%; Score 1821; DB 35; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.3e-180;
```

Db 241 YLVNKPSEERVREIIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLLLVELGFS 300
QY 301 KVFQENPDPFMEINISLEGKTNFFEKRVSEYQRFVMAETTDNVFTLDADF 351
Db 301 KVFQENPDPFMEINISLEGKTNFFEKRVSEYQRFVMAETTDNVFTLDADF 351
RESULT 7
US-60-449-629-94
; Sequence 94, Application US/60449629
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001449
; CURRENT APPLICATION NUMBER: US/60/449, 629
; CURRENT FILING DATE: 2003-02-04
; NUMBER OF SEQ ID NOS: 2712
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 94
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-449-629-94

Query Match 100.0%; Score 1821; DB 37; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.3e-180;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPIWKMYKQAO 60
Db 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPIWKMYKQAO 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAPFAASDGIWNLVERFSQEVQVPEA 120
Db 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAPFAASDGIWNLVERFSQEVQVPEA 120
QY 121 RCFYGOILLIENHSEMYSLIDTYIRDPKREFLFNAIETMPYVKKKADWALRWIADRK 180
Db 121 RCFYGOILLIENHSEMYSLIDTYIRDPKREFLFNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDFACLMFQ 240
Db 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDFACLMFQ 240
QY 241 YLVNKPSEERVREIIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLLLVELGFS 300
Db 241 YLVNKPSEERVREIIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLLLVELGFS 300
QY 301 KVFQENPDPFMEINISLEGKTNFFEKRVSEYQRFVMAETTDNVFTLDADF 351
Db 301 KVFQENPDPFMEINISLEGKTNFFEKRVSEYQRFVMAETTDNVFTLDADF 351

RESULT 8
US-60-505-218-322
; Sequence 322, Application US/60505218
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001482
; CURRENT APPLICATION NUMBER: US/60/505, 218
; CURRENT FILING DATE: 2003-09-24
; NUMBER OF SEQ ID NOS: 22507
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 322
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-505-218-322

Query Match 100.0%; Score 1821; DB 37; Length 351;

Best Local Similarity 100.0%; Pred. No. 1.3e-180;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPIWKMYKQAO 60
Db 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPIWKMYKQAO 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAPFAASDGIWNLVERFSQEVQVPEA 120
Db 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAPFAASDGIWNLVERFSQEVQVPEA 120
QY 121 RCFYGOILLIENHSEMYSLIDTYIRDPKREFLFNAIETMPYVKKKADWALRWIADRK 180
Db 121 RCFYGOILLIENHSEMYSLIDTYIRDPKREFLFNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDFACLMFQ 240
Db 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDFACLMFQ 240
QY 241 YLVNKPSEERVREIIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLLLVELGFS 300
Db 241 YLVNKPSEERVREIIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLLLVELGFS 300
QY 301 KVFQENPDPFMEINISLEGKTNFFEKRVSEYQRFVMAETTDNVFTLDADF 351
Db 301 KVFQENPDPFMEINISLEGKTNFFEKRVSEYQRFVMAETTDNVFTLDADF 351
RESULT 9
US-09-791-537-27671
; Sequence 27671, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27671
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-27671

Query Match 100.0%; Score 1821; DB 22; Length 366;
Best Local Similarity 100.0%; Pred. No. 1.4e-180;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPIWKMYKQAO 60
Db 16 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPIWKMYKQAO 75
QY 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAPFAASDGIWNLVERFSQEVQVPEA 120
Db 76 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAPFAASDGIWNLVERFSQEVQVPEA 135
QY 121 RCFYGOILLIENHSEMYSLIDTYIRDPKREFLFNAIETMPYVKKKADWALRWIADRK 180
Db 136 RCFYGOILLIENHSEMYSLIDTYIRDPKREFLFNAIETMPYVKKKADWALRWIADRK 195
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDFACLMFQ 240
Db 196 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDFACLMFQ 255
QY 241 YLVNKPSEERVREIIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLLLVELGFS 300
Db 256 YLVNKPSEERVREIIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLLLVELGFS 315
QY 301 KVFQENPDPFMEINISLEGKTNFFEKRVSEYQRFVMAETTDNVFTLDADF 351

Db 316 KVFQAEPPDFMENISLEGKTNFFKRVSEYQRFVMAETTDNVFTLDADF 366
|||||

RESULT 10

US-09-629-469A-12811

; Sequence 12811, Application US/09629469A

; GENERAL INFORMATION:

; APPLICANT: OTA, TOSHIO

; APPLICANT: NISHIKAWA, TETSUO

; APPLICANT: HAYASHI, KOJI

; APPLICANT: SAITO, KAORU

; APPLICANT: YAMAMOTO, JUNICHI

; APPLICANT: ISHII, SHIZUKO

; APPLICANT: SUGIYAMA, TOMOYASU

; APPLICANT: WAKAMATSU, AI

; APPLICANT: OTSUKI, TETSUJI

; TITLE OF INVENTION: PRIMERS

; FILE REFERENCE: 084335/0123

; CURRENT APPLICATION NUMBER: US/09/629,469A

; CURRENT FILING DATE: 2000-07-28

; PRIOR APPLICATION NUMBER: JP 1999-248036

; PRIOR FILING DATE: 1999-07-29

; PRIOR APPLICATION NUMBER: JP 1999-300253

; PRIOR FILING DATE: 1999-08-27

; PRIOR APPLICATION NUMBER: JP 2000-118776

; PRIOR FILING DATE: 2000-01-11

; PRIOR APPLICATION NUMBER: JP 2000-183767

; PRIOR FILING DATE: 2000-05-02

; PRIOR APPLICATION NUMBER: JP 2000-241899

; PRIOR FILING DATE: 2000-06-09

; PRIOR APPLICATION NUMBER: 60/159,590

; PRIOR FILING DATE: 1999-10-18

; PRIOR APPLICATION NUMBER: 60/183,322

; PRIOR FILING DATE: 2000-02-17

; NUMBER OF SEQ ID NOS: 19025

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 12811

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-629-469A-12811

Query Match

Best Local Similarity 99.8%; Score 1817; DB 20; Length 351;

Matches 350; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDODERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWKMYKQQAQ 60
Db 1 MGDPERPEAAGLDODERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWKMYKQQAQ 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADKEYFISHILAPFAASDGI VNNLVERFSQEVQVPEA 120
Db 61 ASFWTAEVDLSKDLPHWNKLKADKEYFISHILAPFAASDGI VNNLVERFSQEVQVPEA 120
QY 121 RCFYGFQILLIENHVSEMYSLIDTYIRDPKKREFLNAETMPYVKKKADWALRWIADRK 180
Db 121 RCFYGFQILLIENHVSEMYSLIDTYIRDPKKREFLNAETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPLGTFSNELISRDEGLHCDFACLMFQ 240
Db 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPLGTFSNELISRDEGLHCDFACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIQEFLTEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
Db 241 YLVNKPSEERVREIIVDAVKIQEFLTEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
QY 301 KVFQAEPPDFMENISLEGKTNFFKRVSEYQRFVMAETTDNVFTLDADF 351
Db 301 KVFQAEPPDFMENISLEGKTNFFKRVSEYQRFVMAETTDNVFTLDADF 351

RESULT 11

US-09-791-537-112185

; Sequence 112185, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Bionomix, Inc.

; APPLICANT: Danzer, Joseph

; APPLICANT: Debe, Derek

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS

; TITLE OF INVENTION: METHODS OF USE THEREOF

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: Patent In version 3.0

; SEQ ID NO 112185

; LENGTH: 351

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-791-537-112185

Query Match

Best Local Similarity 99.8%; Score 1817; DB 22; Length 351;

Matches 350; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDODERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWKMYKQQAQ 60
Db 1 MGDPERPEAAGLDODERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWKMYKQQAQ 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADKEYFISHILAPFAASDGI VNNLVERFSQEVQVPEA 120
Db 61 ASFWTAEVDLSKDLPHWNKLKADKEYFISHILAPFAASDGI VNNLVERFSQEVQVPEA 120
QY 121 RCFYGFQILLIENHVSEMYSLIDTYIRDPKKREFLNAETMPYVKKKADWALRWIADRK 180
Db 121 RCFYGFQILLIENHVSEMYSLIDTYIRDPKKREFLNAETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPLGTFSNELISRDEGLHCDFACLMFQ 240
Db 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPLGTFSNELISRDEGLHCDFACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIQEFLTEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
Db 241 YLVNKPSEERVREIIVDAVKIQEFLTEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
QY 301 KVFQAEPPDFMENISLEGKTNFFKRVSEYQRFVMAETTDNVFTLDADF 351
Db 301 KVFQAEPPDFMENISLEGKTNFFKRVSEYQRFVMAETTDNVFTLDADF 351

RESULT 12

US-10-917-503-12811

; Sequence 12811, Application US/10917503

; GENERAL INFORMATION:

; APPLICANT: OTA, TOSHIO

; APPLICANT: NISHIKAWA, TETSUO

; APPLICANT: HAYASHI, KOJI

; APPLICANT: SAITO, KAORU

; APPLICANT: YAMAMOTO, JUNICHI

; APPLICANT: ISHII, SHIZUKO

; APPLICANT: SUGIYAMA, TOMOYASU

; APPLICANT: WAKAMATSU, AI

; APPLICANT: NAGAI, KEIICHI

; APPLICANT: OTSUKI, TETSUJI

; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE

; FILE REFERENCE: 084335/0123

; CURRENT APPLICATION NUMBER: US/10/917,503

; CURRENT FILING DATE: 2004-08-13

; PRIOR APPLICATION NUMBER: US/09/629,469

; PRIOR FILING DATE: 2000-07-28

; PRIOR APPLICATION NUMBER: JP 1999-248036

; PRIOR FILING DATE: 1999-07-29


```

; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12811
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-917-503-12811

Query Match      99.8%; Score 1817; DB 35; Length 351;
Best Local Similarity 99.7%; Pred. No. 3.4e-180;
Matches 350; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPPERPEAAGLDQDERSSDTNESEIKNEEPLLRKSSRRFVIFPIQYPDIMWKYKQAA 60
DB 1 MGDPPERPEAAGLDQDERSSDTNESEIKNEEPLLRKSSRRFVIFPIQYPDIMWKYKQAA 60

QY 61 ASFWTAEEVDLSKDLPHWNKLKADEKFIYSHILAFPAASDGIWVNLVERFSQEVQVPEA 120
DB 61 ASFWTAEEVDLSKDLPHWNKLKADEKFIYSHILAFPAASDGIWVNLVERFSQEVQVPEA 120

QY 121 RCFYGFQIILINVENHSEMYSLIDTYIRDPKREFLNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFYGFQIILINVENHSEMYSLIDTYIRDPKREFLNAIETMPYVKKKADWALRWIADRK 180

QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPLGTFNSNELISRDEGLHCDFACLMFQ 240
DB 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPLGTFNSNELISRDEGLHCDFACLMFQ 240

QY 241 YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLLVGLGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLLVGLGFS 300

QY 301 KVFQAEPPDFMENISLEGKTNFFEKRVSEYQRFVAFVMAETTDNVFTLDADF 351
DB 301 KVFQAEPPDFMENISLEGKTNFFEKRVSEYQRFVAFVMAETTDNVFTLDADF 351

RESULT 13
PCT-US02-10824-143
; Sequence 143, Application PC/TUS0210824
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies
; TITLE OF INVENTION: Prostate Cancer Expression Profiles
; FILE REFERENCE: 9U 206 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/10824
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 60/281,732
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/281,731
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 143
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-10824-143

Query Match      79.8%; Score 1454; DB 1; Length 389;
Best Local Similarity 83.8%; Pred. No. 3.5e-142;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY 31 EEPLLRKSRRFVIFPIQYPDIMWKYKQAAASFWTAEEVDLSKDLPHWNKLKADEKFIY 90
DB 69 DEPLLRENPRFVIFPIEYHDIWQYKKAASFWTAEEVDLSKDIQHWESLKPEERYFIS 128

QY 91 HILAFPAASDGIWVNLVERFSQEVQVPEARCFYGFQIILINVENHSEMYSLIDTYIRDPK 150
DB 129 HVLAFPAASDGIWVNLVERFSQEVQVTEARCFYGFQIAMENIHSEMYSLIDTYIKDPK 188

QY 151 KREFLFNAIETMPYVKKKADWALRWIADRKSTFGERVVAFVMAEVSFSGSFAAIFWLKK 210
DB 189 EREFLFNAIETMPYVKKKADWALRWIGDKEATYGERVVFVAFVMAEVSFSGSFAAIFWLKK 248

QY 211 RGLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEERVREIIVDAVKIEQEFTEAL 270
DB 249 RGLMPGLTFSNELISRDEGLHCDFACLMFKHLVHKPSEERVREIIVDAVKIEQEFTEAL 308

QY 271 PVGLIGMNCILMKQYIEFVADRLLVGLGFSKVFOAENPFDPMENISLEGKTNFFEKRVSE 330
DB 309 PVKLGIMNCILMKQYIEFVADRLMLGLGFSKVFRVENPFDPMENISLEGKTNFFEKRVSE 368

RESULT 14
PCT-US03-17409-171
; Sequence 171, Application PC/TUS0317409
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION
; APPLICANT: RICKERT, Paula K.
; APPLICANT: KEASNOW, Randi
; TITLE OF INVENTION: DIAGNOSTIC MARKERS FOR LUNG CANCER
; FILE REFERENCE: PA-0051 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/17409
; CURRENT FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: US 60/386,005
; PRIOR FILING DATE: 2002-06-04
; NUMBER OF SEQ ID NOS: 296
; SOFTWARE: PERL Program
; SEQ ID NO 171
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 000899CD1
PCT-US03-17409-171

Query Match      79.8%; Score 1454; DB 1; Length 389;
Best Local Similarity 83.8%; Pred. No. 3.5e-142;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY 31 EEPLLRKSRRFVIFPIQYPDIMWKYKQAAASFWTAEEVDLSKDLPHWNKLKADEKFIY 90
DB 69 DEPLLRENPRFVIFPIEYHDIWQYKKAASFWTAEEVDLSKDIQHWESLKPEERYFIS 128

QY 91 HILAFPAASDGIWVNLVERFSQEVQVPEARCFYGFQIILINVENHSEMYSLIDTYIRDPK 150
DB 129 HVLAFPAASDGIWVNLVERFSQEVQVTEARCFYGFQIAMENIHSEMYSLIDTYIKDPK 188

QY 151 KREFLFNAIETMPYVKKKADWALRWIADRKSTFGERVVAFVMAEVSFSGSFAAIFWLKK 210
DB 189 EREFLFNAIETMPYVKKKADWALRWIGDKEATYGERVVFVAFVMAEVSFSGSFAAIFWLKK 248

QY 211 RGLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEERVREIIVDAVKIEQEFTEAL 270
DB 249 RGLMPGLTFSNELISRDEGLHCDFACLMFKHLVHKPSEERVREIIVDAVKIEQEFTEAL 308

QY 271 PVGLIGMNCILMKQYIEFVADRLLVGLGFSKVFOAENPFDPMENISLEGKTNFFEKRVSE 330
DB 309 PVKLGIMNCILMKQYIEFVADRLMLGLGFSKVFRVENPFDPMENISLEGKTNFFEKRVSE 368
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2005, 06:56:26 ; Search time 44 Seconds
(without alignments)
1802.648 Million cell updates/sec

Title: US-10-698-228-1

Perfect score: 1821

Sequence: 1 MGDPERPEAGLDQDSSSS.....QRFVVAETTDNVFTLDADF 351

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 972358 seqs, 225972999 residues

Total number of hits satisfying chosen parameters: 972358

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA New.*

- 1: /cgn2_6/prodata/1/paa/PCT_NEW_COMB.pep.*
- 2: /cgn2_6/prodata/1/paa/US05_NEW_COMB.pep.*
- 3: /cgn2_6/prodata/1/paa/US07_NEW_COMB.pep.*
- 4: /cgn2_6/prodata/1/paa/US08_NEW_COMB.pep.*
- 5: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep.*
- 6: /cgn2_6/prodata/1/paa/US10_NEW_COMB.pep.*
- 7: /cgn2_6/prodata/1/paa/US11_NEW_COMB.pep.*
- 8: /cgn2_6/prodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	100.0	351	1	PCT-US05-09639-15
2	1821	100.0	351	1	PCT-US05-21650-34
3	1821	100.0	351	6	US-10-990-328A-7547
4	1821	100.0	351	7	US-11-088-686-15
5	1817	99.8	351	1	PCT-US05-09639-17
6	1817	99.8	351	7	US-11-088-686-17
7	1816	99.7	351	1	PCT-US05-09639-19
8	1816	99.7	351	7	US-11-088-686-19
9	1813	99.6	351	1	PCT-US05-09639-21
10	1813	99.6	351	7	US-11-088-686-21
11	1812	99.5	351	1	PCT-US05-09639-23
12	1812	99.5	351	1	PCT-US05-09639-25
13	1812	99.5	351	7	US-11-088-686-23
14	1812	99.5	351	7	US-11-088-686-25
15	1811	99.5	351	1	PCT-US05-09639-27
16	1811	99.5	351	7	US-11-088-686-27
17	1454	79.8	389	1	PCT-US03-10870-559
18	1454	79.8	389	1	PCT-US05-09639-1
19	1454	79.8	389	1	PCT-US05-21650-33
20	1454	79.8	389	6	US-10-940-774A-6016
21	1454	79.8	389	7	US-11-088-686-1
22	1454	79.8	449	6	US-10-990-328A-12349
23	1454	79.8	449	6	US-10-990-328A-12350
24	1454	79.8	453	6	US-10-450-763-45646
25	1454	79.8	453	6	US-10-940-774A-7896

26	1450	79.6	389	1	PCT-US05-09639-3	Sequence 3, Appli
27	1450	79.6	389	7	US-11-088-686-3	Sequence 3, Appli
28	1449	79.6	389	1	PCT-US05-09639-5	Sequence 5, Appli
29	1449	79.6	389	7	US-11-088-686-5	Sequence 5, Appli
30	1446	79.4	389	1	PCT-US05-09639-7	Sequence 7, Appli
31	1446	79.4	389	7	US-11-088-686-7	Sequence 7, Appli
32	1445	79.4	389	1	PCT-US05-09639-9	Sequence 9, Appli
33	1445	79.4	389	7	US-11-088-686-9	Sequence 9, Appli
34	1445	79.4	389	1	PCT-US05-09639-11	Sequence 11, Appl
35	1445	79.4	389	7	US-11-088-686-11	Sequence 11, Appl
36	1444	79.3	389	1	PCT-US05-09639-13	Sequence 13, Appl
37	1444	79.3	389	7	US-11-088-686-13	Sequence 13, Appl
38	1371	75.3	389	8	US-60-669-241-28508	Sequence 28508, A
39	1348	74.0	393	7	US-11-097-143-19182	Sequence 19182, A
40	1274	70.0	319	1	PCT-US05-12947-102	Sequence 102, App
41	1232.5	67.7	404	6	US-10-703-032-123588	Sequence 123588,
42	1156	63.5	330	7	US-11-251-208-421	Sequence 421, App
43	1147.5	63.0	329	7	US-11-251-208-427	Sequence 427, App
44	1128.5	62.0	327	7	US/11/241	GENERAL INFORMATI
45	1125.5	61.8	345	7	US-11-251-208-519	Sequence 519, App

ALIGNMENTS

RESULT 1
PCT-US05-09639-15
; Sequence 15, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004WO1
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-15

Query Match	100.0%;	Score 1821;	DB 1;	Length 351;
Best Local Similarity	100.0%;	Pred. No. 3.3e-160;		
Matches 351;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MGDPERPEAGLDQDSSSDTNESEIKSNEEPLLRKSSRRFVFPQYPDIMWKYKQAQ	60	
Db	1	MGDPERPEAGLDQDSSSDTNESEIKSNEEPLLRKSSRRFVFPQYPDIMWKYKQAQ	60	
Qy	61	ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAFFAASDGI VNNELVERFSQEVQVPEA	120	
Db	61	ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAFFAASDGI VNNELVERFSQEVQVPEA	120	
Qy	121	RCFYGFQILNIENVHSEMYSLIDITYIRDPKRELFNAIETMPYVKKKADWALRIADRK	180	
Db	121	RCFYGFQILNIENVHSEMYSLIDITYIRDPKRELFNAIETMPYVKKKADWALRIADRK	180	
Qy	181	STGERVVAFAAEGVEGFFSGSFAAI FWLKKRGLNPGLTFSNELISRDEGLHCDFACLMFQ	240	
Db	181	STGERVVAFAAEGVEGFFSGSFAAI FWLKKRGLNPGLTFSNELISRDEGLHCDFACLMFQ	240	
Qy	241	YLNKKSSEERVREIIVDAVKIEQBELTEALPVGLIGNCILMKOYIFVADRLLVELGFS	300	
Db	241	YLNKKSSEERVREIIVDAVKIEQBELTEALPVGLIGNCILMKOYIFVADRLLVELGFS	300	
Qy	301	KVFQAEVNFDFMENISLEGKTNFFEKRVSEYQRFVAVNAETTDNVFTLDADF	351	
Db	301	KVFQAEVNFDFMENISLEGKTNFFEKRVSEYQRFVAVNAETTDNVFTLDADF	351	

```
RESULT 2
PCT-US05-21650-34
; Sequence 34, Application PC/TUS0521650
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: MIGFS AS MODIFIERS OF THE IGF PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX05-017
; CURRENT APPLICATION NUMBER: PCT/US05/21650
; CURRENT FILING DATE: 2005-06-27
; PRIOR APPLICATION NUMBER: 60/581,689
; PRIOR FILING DATE: 2004-06-21
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 34
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-21650-34

Query Match      100.0%; Score 1821; DB 1; Length 351;
Best Local Similarity 100.0%; Pred. No. 3.3e-160;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPFIQYDPDIWKMYKQAA 60
DB 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPFIQYDPDIWKMYKQAA 60
QY 61 ASFWTAAEEVDLSKDLPHWNKLKADEKYFISHILAFPAASDGI VNNELVERFSQEVQVPEA 120
DB 61 ASFWTAAEEVDLSKDLPHWNKLKADEKYFISHILAFPAASDGI VNNELVERFSQEVQVPEA 120
QY 121 RCFYGFQI LIENVHSEMYSLIDTYIRDPKREFL FNAIETMPYVKKADWALRWIADRK 180
DB 121 RCFYGFQI LIENVHSEMYSLIDTYIRDPKREFL FNAIETMPYVKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDPACLMFQ 240
DB 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
QY 301 KVFQAE NPPDFMENISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 301 KVFQAE NPPDFMENISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 3
US-10-990-328A-7547
; Sequence 7547, Application US/10990328A
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328A
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7547
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328A-7547

Query Match      100.0%; Score 1821; DB 6; Length 351;
Best Local Similarity 100.0%; Pred. No. 3.3e-160;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPFIQYDPDIWKMYKQAA 60
DB 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPFIQYDPDIWKMYKQAA 60
QY 61 ASFWTAAEEVDLSKDLPHWNKLKADEKYFISHILAFPAASDGI VNNELVERFSQEVQVPEA 120
DB 61 ASFWTAAEEVDLSKDLPHWNKLKADEKYFISHILAFPAASDGI VNNELVERFSQEVQVPEA 120
QY 121 RCFYGFQI LIENVHSEMYSLIDTYIRDPKREFL FNAIETMPYVKKADWALRWIADRK 180
DB 121 RCFYGFQI LIENVHSEMYSLIDTYIRDPKREFL FNAIETMPYVKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDPACLMFQ 240
DB 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
QY 301 KVFQAE NPPDFMENISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 301 KVFQAE NPPDFMENISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 4
US-11-088-686-15
; Sequence 15, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-15

Query Match      100.0%; Score 1821; DB 7; Length 351;
Best Local Similarity 100.0%; Pred. No. 3.3e-160;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPFIQYDPDIWKMYKQAA 60
DB 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPFIQYDPDIWKMYKQAA 60
QY 61 ASFWTAAEEVDLSKDLPHWNKLKADEKYFISHILAFPAASDGI VNNELVERFSQEVQVPEA 120
DB 61 ASFWTAAEEVDLSKDLPHWNKLKADEKYFISHILAFPAASDGI VNNELVERFSQEVQVPEA 120
QY 121 RCFYGFQI LIENVHSEMYSLIDTYIRDPKREFL FNAIETMPYVKKADWALRWIADRK 180
DB 121 RCFYGFQI LIENVHSEMYSLIDTYIRDPKREFL FNAIETMPYVKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDPACLMFQ 240
DB 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
QY 301 KVFQAE NPPDFMENISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 301 KVFQAE NPPDFMENISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 5
PCT-US05-09639-17
```

```
; Sequence 17, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-17

Query Match          99.8%; Score 1817; DB 1; Length 351;
Best Local Similarity 99.7%; Pred. No. 7.9e-160;
Matches 350; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDSSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWMYKQQAQ 60
DB 1 MGDPERPEAAGLDQDSSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWMYKQQAQ 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAPFAASDGI VNNLVVERFSQEVQVPEA 120
DB 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAPFAASDGI VNNLVVERFSQEVQVPEA 120
QY 121 RCFYGFQIILIENHSEMYSLIDITYIRDPKRELFNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFYGFQIILIENHSEMYSLIDITYIRDPKRELFNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGKGLMPGLTFSNELISRDEGLHCDPACLMFQ 240
DB 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGKGLMPGLTFSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
QY 301 KVFQAEPPDFMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTILDADF 351
DB 301 KVFQAEPPDFMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTILDADF 351

RESULT 7
PCT-US05-09639-19
; Sequence 19, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-19

Query Match          99.7%; Score 1816; DB 1; Length 351;
Best Local Similarity 99.7%; Pred. No. 9.7e-160;
Matches 350; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDSSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWMYKQQAQ 60
DB 1 MGDPERPEAAGLDQDSSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWMYKQQAQ 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAPFAASDGI VNNLVVERFSQEVQVPEA 120
DB 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAPFAASDGI VNNLVVERFSQEVQVPEA 120
QY 121 RCFYGFQIILIENHSEMYSLIDITYIRDPKRELFNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFYGFQIILIENHSEMYSLIDITYIRDPKRELFNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGKGLMPGLTFSNELISRDEGLHCDPACLMFQ 240
DB 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGKGLMPGLTFSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
QY 301 KVFQAEPPDFMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTILDADF 351
DB 301 KVFQAEPPDFMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTILDADF 351

RESULT 8
US-11-088-686-19
; Sequence 19, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: US/11/088,686
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-17

Query Match          99.8%; Score 1817; DB 7; Length 351;
Best Local Similarity 99.7%; Pred. No. 7.9e-160;
Matches 350; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDSSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWMYKQQAQ 60
DB 1 MGDPERPEAAGLDQDSSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWMYKQQAQ 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAPFAASDGI VNNLVVERFSQEVQVPEA 120
DB 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAPFAASDGI VNNLVVERFSQEVQVPEA 120
QY 121 RCFYGFQIILIENHSEMYSLIDITYIRDPKRELFNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFYGFQIILIENHSEMYSLIDITYIRDPKRELFNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGKGLMPGLTFSNELISRDEGLHCDPACLMFQ 240
DB 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGKGLMPGLTFSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
QY 301 KVFQAEPPDFMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTILDADF 351
DB 301 KVFQAEPPDFMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTILDADF 351
```

```
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-19

Query Match      99.7%; Score 1816; DB 7; Length 351;
Best Local Similarity 99.7%; Pred. No. 9.7e-160;
Matches 350; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDODERSSDTNESEIKNEEPLLRKSSRRFVIFPIQYDPDIWMYKQAO 60
DB 1 MGDPERPEAAGLDODERSSDTNESEIKNEEPLLRKSSRRFVIFPIQYDPDIWMYKQAO 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAFFAASDGI VNNLVVERFSQEVQVPEA 120
DB 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAFFAASDGI VNNLVVERFSQEVQVPEA 120
QY 121 RCFYGFQILLENVHSEMYSLIDITYIRDPKKREFLNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFYGFQILLENVHSEMYSLIDITYIRDPKKREFLNAIETMPYVKKKADWALRWIADRK 180
QY 121 RCFWGFQILLENVHSEMYSLIDITYIRDPKKREFLNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFWGFQILLENVHSEMYSLIDITYIRDPKKREFLNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDPACLMFQ 240
DB 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGNCILMKQYIEFVADRLLVELGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGNCILMKQYIEFVADRLLVELGFS 300
QY 301 KVFQENPPDFMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 301 KVFQENPPDFMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 9
PCT-US05-09639-21
; Sequence 21, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-21

Query Match      99.6%; Score 1813; DB 1; Length 351;
Best Local Similarity 99.4%; Pred. No. 1.18e-159;
Matches 349; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDODERSSDTNESEIKNEEPLLRKSSRRFVIFPIQYDPDIWMYKQAO 60
DB 1 MGDPERPEAAGLDODERSSDTNESEIKNEEPLLRKSSRRFVIFPIQYDPDIWMYKQAO 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAFFAASDGI VNNLVVERFSQEVQVPEA 120
DB 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAFFAASDGI VNNLVVERFSQEVQVPEA 120
QY 121 RCFYGFQILLENVHSEMYSLIDITYIRDPKKREFLNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFYGFQILLENVHSEMYSLIDITYIRDPKKREFLNAIETMPYVKKKADWALRWIADRK 180
QY 121 RCFWGFQILLENVHSEMYSLIDITYIRDPKKREFLNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFWGFQILLENVHSEMYSLIDITYIRDPKKREFLNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDPACLMFQ 240
DB 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGNCILMKQYIEFVADRLLVELGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGNCILMKQYIEFVADRLLVELGFS 300
QY 301 KVFQENPPDFMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 301 KVFQENPPDFMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 10
US-11-088-686-21
; Sequence 21, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-21

Query Match      99.6%; Score 1813; DB 7; Length 351;
Best Local Similarity 99.4%; Pred. No. 1.8e-159;
Matches 349; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDODERSSDTNESEIKNEEPLLRKSSRRFVIFPIQYDPDIWMYKQAO 60
DB 1 MGDPERPEAAGLDODERSSDTNESEIKNEEPLLRKSSRRFVIFPIQYDPDIWMYKQAO 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAFFAASDGI VNNLVVERFSQEVQVPEA 120
DB 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAFFAASDGI VNNLVVERFSQEVQVPEA 120
QY 121 RCFYGFQILLENVHSEMYSLIDITYIRDPKKREFLNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFYGFQILLENVHSEMYSLIDITYIRDPKKREFLNAIETMPYVKKKADWALRWIADRK 180
QY 121 RCFWGFQILLENVHSEMYSLIDITYIRDPKKREFLNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFWGFQILLENVHSEMYSLIDITYIRDPKKREFLNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDPACLMFQ 240
DB 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGNCILMKQYIEFVADRLLVELGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGNCILMKQYIEFVADRLLVELGFS 300
QY 301 KVFQENPPDFMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 301 KVFQENPPDFMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 11
PCT-US05-09639-23
; Sequence 23, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
```



```
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-23

Query Match          99.5%; Score 1812; DB 1; Length 351;
Best Local Similarity 99.4%; Pred. No. 2.3e-159;
Matches 349; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPPIQYPDIMWKYKQAO 60
DB 1 MGDPERPAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPPIQYPDIMWKYKQAO 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADKEYFISHILAFPAASDGI VNNLVERFSQEVQVPEA 120
DB 61 ASFWTAEVDLSKDLPHWNKLKADKEYFISHILAFPAASDGI VNNLVERFSQEVQVPEA 120
QY 121 RCFYGFQILLIENHSEMYSLIDTYIRDPKKEFLFNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFYGFQILLIENHSEMYSLIDTYIRDPKKEFLFNAIETMPYVKKKADWALRWIADRK 180
QY 121 RCFYGFQILLIENHSEMYSLIDTYIRDPKKEFLFNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFYGFQILLIENHSEMYSLIDTYIRDPKKEFLFNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFPAAVEGVFFSGSFAAIFWLKRGMLMPLGTFSNELISRDEGLHCDPACLMFQ 240
DB 181 STFGERVVAFPAAVEGVFFSGSFAAIFWLKRGMLMPLGTFSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVVELGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVVELGFS 300
QY 301 KVFQENPFDPMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 301 KVFQENPFDPMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 12
PCT-US05-09639-25
; Sequence 25, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-25

Query Match          99.5%; Score 1812; DB 1; Length 351;
Best Local Similarity 99.4%; Pred. No. 2.3e-159;
Matches 349; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPPIQYPDIMWKYKQAO 60
DB 1 MGDPERPAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPPIQYPDIMWKYKQAO 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADKEYFISHILAFPAASDGI VNNLVERFSQEVQVPEA 120
DB 61 ASFWTAEVDLSKDLPHWNKLKADKEYFISHILAFPAASDGI VNNLVERFSQEVQVPEA 120
QY 121 RCFYGFQILLIENHSEMYSLIDTYIRDPKKEFLFNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFYGFQILLIENHSEMYSLIDTYIRDPKKEFLFNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFPAAVEGVFFSGSFAAIFWLKRGMLMPLGTFSNELISRDEGLHCDPACLMFQ 240
DB 181 STFGERVVAFPAAVEGVFFSGSFAAIFWLKRGMLMPLGTFSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVVELGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVVELGFS 300
QY 301 KVFQENPFDPMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 301 KVFQENPFDPMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 13
US-11-088-686-23
; Sequence 23, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-23

Query Match          99.5%; Score 1812; DB 7; Length 351;
Best Local Similarity 99.4%; Pred. No. 2.3e-159;
Matches 349; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPPIQYPDIMWKYKQAO 60
DB 1 MGDPERPAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPPIQYPDIMWKYKQAO 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADKEYFISHILAFPAASDGI VNNLVERFSQEVQVPEA 120
DB 61 ASFWTAEVDLSKDLPHWNKLKADKEYFISHILAFPAASDGI VNNLVERFSQEVQVPEA 120
QY 121 RCFYGFQILLIENHSEMYSLIDTYIRDPKKEFLFNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFYGFQILLIENHSEMYSLIDTYIRDPKKEFLFNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFPAAVEGVFFSGSFAAIFWLKRGMLMPLGTFSNELISRDEGLHCDPACLMFQ 240
DB 181 STFGERVVAFPAAVEGVFFSGSFAAIFWLKRGMLMPLGTFSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVVELGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVVELGFS 300
QY 301 KVFQENPFDPMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 301 KVFQENPFDPMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 14
US-11-088-686-25
; Sequence 25, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
```

;
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-25

Query Match
Best Local Similarity 99.4%; Score 1812; DB 7; Length 351;
Matches 349; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGDPERPEAAGLDQDERSSDTNESIKNEEPLLRKSSRRFVIFPIQYDPIWKYKQAO 60
Db 1 MGDPERPEAAGLDQDERSSDTNESIKNEEPLLRKSSRRFVIFPIQYDPIWKYKQAO 60
QY 61 ASFWTAAEVDLSKOLPHWNKLKADEKYFISHILAPFAASDGI VNNLVERFSOEVOVPEA 120
Db 61 ASFWTAAEVDLSKOLPHWNKLKADEKYFISHILAPFAASDGI VNNLVERFSOEVOVPEA 120
QY 121 RCFYGFQILINVENHSEMSLLIDTYIRDPKREFLNAIETMPYVKKADWALRWIADRK 180
Db 121 RCFYGFQILINVENHSEMSLLIDTYIRDPKREFLNAIETMPYVKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAI FWLKKRGLMPLGTFSNELISRDEGLHCDPACLMFQ 240
Db 181 STFGERVVAFAAVEGVFFSGSFAAI FWLKKRGLMPLGTFSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLLVELGFS 300
Db 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLLVELGFS 300
QY 301 KVFQAEENPDPFMEINISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF 351
Db 301 KVFQAEENPDPFMEINISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 15
PCT-US05-09639-27
; Sequence 27, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-27

Query Match
Best Local Similarity 99.4%; Score 1811; DB 1; Length 351;
Matches 349; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGDPERPEAAGLDQDERSSDTNESIKNEEPLLRKSSRRFVIFPIQYDPIWKYKQAO 60
Db 1 MGDPERPEAAGLDQDERSSDTNESIKNEEPLLRKSSRRFVIFPIQYDPIWKYKQAO 60
QY 61 ASFWTAAEVDLSKOLPHWNKLKADEKYFISHILAPFAASDGI VNNLVERFSOEVOVPEA 120
Db 61 ASFWTAAEVDLSKOLPHWNKLKADEKYFISHILAPFAASDGI VNNLVERFSOEVOVPEA 120
QY 121 RCFYGFQILINVENHSEMSLLIDTYIRDPKREFLNAIETMPYVKKADWALRWIADRK 180
Db 121 RCFYGFQILINVENHSEMSLLIDTYIRDPKREFLNAIETMPYVKKADWALRWIADRK 180

Db 121 RCFYGFQILINVENHSEMSLLIDTYIRDPKREFLNAIETMPYVKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAI FWLKKRGLMPLGTFSNELISRDEGLHCDPACLMFQ 240
Db 181 STFGERVVAFAAVEGVFFSGSFAAI FWLKKRGLMPLGTFSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLLVELGFS 300
Db 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLLVELGFS 300
QY 301 KVFQAEENPDPFMEINISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF 351
Db 301 KVFQAEENPDPFMEINISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF 351

Search completed: October 30, 2005, 07:02:43
Job time : 45 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 30, 2005, 06:24:21 ; Search time 173.25 Seconds
(without alignments)
14198.102 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1846

Sequence: 1 atggcgaccgaggaaggcc.....tcaccttgatgcagatttt 1053

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 695266 seqs, 1168006243 residues

Total number of hits satisfying chosen parameters: 13918532

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10698228/runat_26102005_100609_4045/app.query.fasta_1.2446
-DB=Pending Patents AA Main -QFMT=fastan -SUFFIX=n2p.rapm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS-human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORW=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10698228@cgn_1_1_523@runat_26102005_100609_4045
-NCPU=6 -ICPU=3 -NO MMAP -LARGEMEMORY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FCGAPOP=6 -FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Pending Patents AA Main.*
2: /cgn2_6/ptodata/1/paa/PTUS.COMB.pcp.*
3: /cgn2_6/ptodata/1/paa/US06.COMB.pcp.*
4: /cgn2_6/ptodata/1/paa/US07.COMB.pcp.*
5: /cgn2_6/ptodata/1/paa/US08.COMB.pcp.*
6: /cgn2_6/ptodata/1/paa/US082.COMB.pcp.*
7: /cgn2_6/ptodata/1/paa/US083.COMB.pcp.*
8: /cgn2_6/ptodata/1/paa/US084.COMB.pcp.*
9: /cgn2_6/ptodata/1/paa/US085.COMB.pcp.*
10: /cgn2_6/ptodata/1/paa/US086.COMB.pcp.*
11: /cgn2_6/ptodata/1/paa/US087.COMB.pcp.*
12: /cgn2_6/ptodata/1/paa/US088.COMB.pcp.*
13: /cgn2_6/ptodata/1/paa/US089.COMB.pcp.*
14: /cgn2_6/ptodata/1/paa/US090.COMB.pcp.*
15: /cgn2_6/ptodata/1/paa/US091.COMB.pcp.*
16: /cgn2_6/ptodata/1/paa/US092.COMB.pcp.*
17: /cgn2_6/ptodata/1/paa/US093.COMB.pcp.*
18: /cgn2_6/ptodata/1/paa/US094.COMB.pcp.*
19: /cgn2_6/ptodata/1/paa/US095.COMB.pcp.*
20: /cgn2_6/ptodata/1/paa/US096.COMB.pcp.*
21: /cgn2_6/ptodata/1/paa/US097A.COMB.pcp.*
22: /cgn2_6/ptodata/1/paa/US097B.COMB.pcp.*
23: /cgn2_6/ptodata/1/paa/US098.COMB.pcp.*
24: /cgn2_6/ptodata/1/paa/US099A.COMB.pcp.*
25: /cgn2_6/ptodata/1/paa/US099B.COMB.pcp.*
26: /cgn2_6/ptodata/1/paa/US100.COMB.pcp.*
27: /cgn2_6/ptodata/1/paa/US101.COMB.pcp.*
28: /cgn2_6/ptodata/1/paa/US102.COMB.pcp.*

29: /cgn2_6/ptodata/1/paa/US103.COMB.pcp.*
30: /cgn2_6/ptodata/1/paa/US104.COMB.pcp.*
31: /cgn2_6/ptodata/1/paa/US105.COMB.pcp.*
32: /cgn2_6/ptodata/1/paa/US106.COMB.pcp.*
33: /cgn2_6/ptodata/1/paa/US107.COMB.pcp.*
34: /cgn2_6/ptodata/1/paa/US108.COMB.pcp.*
35: /cgn2_6/ptodata/1/paa/US109.COMB.pcp.*
36: /cgn2_6/ptodata/1/paa/US110.COMB.pcp.*
37: /cgn2_6/ptodata/1/paa/US60.COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1821	98.6	351	22	US-09-787-491B-11
2	1821	98.6	351	22	US-09-791-537-77474
3	1821	98.6	351	27	US-10-170-205E-19045
4	1821	98.6	351	32	US-10-698-228-1
5	1821	98.6	351	35	US-10-990-328-7547
6	1821	98.6	351	37	US-60-128-660-7
7	1821	98.6	351	37	US-60-449-629-94
8	1821	98.6	351	37	US-60-505-218-322
9	1821	98.6	366	22	US-09-791-537-27671
10	1817	98.4	351	20	US-09-629-469A-12811
11	1817	98.4	351	22	US-09-791-537-112185
12	1817	98.4	351	35	US-10-917-503-12811
13	1454	78.8	389	1	PCT-US02-10824-143
14	1454	78.8	389	1	PCT-US03-17409-171
15	1454	78.8	389	10	US-08-663-617-33
16	1454	78.8	389	12	US-08-837-599-33
17	1454	78.8	389	27	US-09-791-537-88179
18	1454	78.8	389	27	US-10-170-205B-8018
19	1454	78.8	389	30	US-10-408-765-559
20	1454	78.8	389	30	US-10-408-765A-559
21	1454	78.8	389	32	US-10-698-228-5
22	1454	78.8	389	33	US-10-733-878-457
23	1454	78.8	389	33	US-10-756-149-4899
24	1454	78.8	389	35	US-10-940-774-6016
25	1454	78.8	389	37	US-60-389-987-559
26	1454	78.8	389	37	US-60-412-418-559
27	1454	78.8	389	37	US-60-452-680-20616
28	1454	78.8	389	37	US-60-568-073-984
29	1454	78.8	413	1	PCT-US00-05882-1347
30	1454	78.8	413	24	US-09-325-301-1347
31	1454	78.8	449	35	US-10-990-328-12349
32	1454	78.8	449	35	US-10-990-328-12350
33	1454	78.8	453	1	PCT-US01-04926A-686
34	1454	78.8	453	1	PCT-US01-08631-45646
35	1454	78.8	453	28	US-10-220-335-686
36	1454	78.8	453	28	US-10-290-752-686
37	1454	78.8	453	35	US-10-940-774-7896
38	1443	78.2	390	10	US-08-663-617-35
39	1443	78.2	390	12	US-08-837-599-35
40	1443	78.2	390	22	US-09-791-537-117257
41	1425.5	77.2	386	22	US-09-791-537-114288
42	1425.5	77.2	386	30	US-10-403-571-76
43	1425.5	77.2	386	30	US-10-405-887-76
44	1421	77.0	386	22	US-09-791-537-36076
45	1349.5	73.1	399	22	US-09-791-537-1129

ALIGNMENTS

RESULT 1

US-09-787-491B-11
; Sequence 11, Application US/09787491B
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; TANG, Y. Tom;

APPLICANT: CORLEY, Neil C.; GUEGLER, Karl J.;
APPLICANT: GORONE, Gina A.; PATTERSON, Chandra;
APPLICANT: HILLMAN, Jennifer L.; BAUGHN, Mariah R.;
APPLICANT: LAL, Preeti; AZIMZAI, Yalda;
APPLICANT: YUE, Henry; YANG, Junming
TITLE OF INVENTION: RNA-ASSOCIATED PROTEINS
FILE REFERENCE: PF-0600 USN
CURRENT APPLICATION NUMBER: US/09/787,491B
CURRENT FILING DATE: 2002-11-04
PRIOR APPLICATION NUMBER: PCT/US99/21688
PRIOR FILING DATE: 1999-09-17
PRIOR APPLICATION NUMBER: US 60/128,660
PRIOR FILING DATE: 1999-04-08
PRIOR APPLICATION NUMBER: US 60/069,391
PRIOR FILING DATE: 1998-11-04
PRIOR APPLICATION NUMBER: US 60/183,025
PRIOR FILING DATE: 1998-09-22
PRIOR APPLICATION NUMBER: US 60/155,246
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: US 09/158,720
PRIOR FILING DATE: 1998-09-17
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PERL Program
SEQ ID NO 11
LENGTH: 351
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 2604449CD1
US-09-787-491B-11

Alignment Scores:
Pred. No.: 1.89e-186 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 22 Gaps: 0

US-10-698-228-2 (1-1053) x US-09-787-491B-11 (1-351)

```
QY 1 ATGGGCGACCGGAAAGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAAGAAAGTGAATAAGTCAATGAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
DB 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerArg 40
QY 121 CGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAATGTATTAACAGGCACAG 180
DB 41 ArgPheValIlePheProIleGlnTyProAspIleTrpLysMetTyLysGlnAlaGln 60
QY 181 GCTTCTCTCTGACACAGAGAGGTCGACTTATCAAGAGTCTCCTCACTGGAACAAG 240
DB 61 AlaSerPheTrpThrAlaGluGluValAlaAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAAGCAGATGAGAAGTACTCTCATCTCTACATCTTACAGCTTTTTCGAGCCAGTGAT 300
DB 81 LeuLysAlaAspGluLysTyPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTGTAATAAGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCCAGAGGCT 360
DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnProGluAla 120
QY 361 CGCTGTTCTATGCGCTTCAATTTCTCATCGAGAATGTTCACTCAGAGATGTACATTTG 420
DB 121 ArgCysPheTyGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTySerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGAA 480
DB 141 LeuIleAspThrTyrlleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
```

```
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCCTTGGATGATAGCAGATAGAAAA 540
DB 161 ThrMetProTyValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAAGAGTTTCTTCTCAGA 600
DB 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTTCTGCTTAAAGAGAGAGGTCTTATGCCAGGACTCATTCTTTCC 560
DB 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACCTGTGACTTTGCTTGTGCTGCTGATGTTC 720
DB 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATAAGCTTTCAGAAAGAGGTGAGGAGATCATTTGATGCTCTCAAA 780
DB 241 TyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGAGCAGAGCTTTTAAACAGAGCCTTGCCAGTTGCCCTCATTTGGAATGAAATTCATT 840
DB 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAAACAGTACATTGAGTTTGTAGCTCACAGATTACTTGTGGAACCTTGGATTCTCA 900
DB 281 LeuMetLysGlnTyrlleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGGCAGAAAAATCCTTTTGATTTTATGAAAAACATTTCTTTAGAAAGAAAA 960
DB 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTGAGAAACAGATTTCAGAGTATCAGCGTTTTCAGAGTTTATGCGACAAACC 1020
DB 321 ThrAsnPhePheGluLysArgValSerGluTyArgLysArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGTCCTTCACTTTGGATGCGAGATTTT 1053
DB 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 2

US-09-791-537-77474
Sequence 77474, Application US/09791537
GENERAL INFORMATION:

APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 77474
LENGTH: 351
TYPE: PRT
ORGANISM: Homo sapiens
US-09-791-537-77474

Alignment Scores:
Pred. No.: 1.89e-186 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 22 Gaps: 0

US-10-698-228-2 (1-1053) x US-09-791-537-77474 (1-351)

```
QY 1 ATGGGCGACCGGAAAGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
DB 1
```


QY 721 TACTTAGTAATAAGCTTTCAGAGAAAGGGTCAGGAGATCATTTGTCATGCTGTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260
QY 781 ATTGACGAGGAGTTTAAACAGAGCTTGCAGTTGGCTCATTTGGAATGAATTCAT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGCAGAAATCCTTTGATTTATGGAACACATTTCTTTAGAGGAAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluLys 320
QY 961 ACAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGGATTTGCGAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAAGCTCTTACCTTGGATGCAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 4

US-10-698-228-1
; Sequence 1, Application US/10698228
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; PRIOR FILING DATE: 2003-10-30
; PRIOR FILING DATE: 2001-12-28
; PRIOR FILING DATE: 2001-12-28
; PRIOR FILING DATE: 2001-12-28
; PRIOR FILING DATE: 1999-06-28
; PRIOR FILING DATE: 1999-06-28
; PRIOR FILING DATE: 1999-07-06
; PRIOR FILING DATE: 1999-07-06
; PRIOR FILING DATE: 2000-01770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 1
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-1

Alignment Scores:
Pred. No.: 1.89e-186 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 32 Gaps: 0

US-10-698-228-2 (1-1053) x US-10-698-228-1 (1-351)

QY 1 ATGGGCGCCCGAAGGCGGAGCGCGGCTCGATCAGGATCAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaIleGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAAGTAAAGTCAAAATGAAGAGCCACTCCTCAAGAAAGAGTTCTCGC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
QY 121 CGGTTTGTTCATTTTCCATCCAGTACCTGATATTTGGAAATGATTAACAGGACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTTCGACAGACAGAGGTCGACTTATCAAGAGTCTCCCTCACTGGAACAG 240

Db 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTTAAAGCAGATCAGAGTACTTCTCTCAGATCTTAGCTTTTGGACGACGATCAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTCGAGGTTCCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTCAAAATTCATCATCGAGAAATGTTCTCATCAGAGATGTCAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTCATGACACATTTACATCAGATCCCAAGAAAGGCAATTTTATTAATGCAATTCGA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTGGCATGGATAGCAGATAGAAAA 540
Db 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTGTGCTGCTAGAGAGTCTTCTTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaValGluGluValPhePheSerGly 200
QY 601 TCTTTTGTCTATATCTGGCTTAAAGAGAGAGTCTTATGCCAGGACTCACATTTTCC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAATCATCAGCAGATGAAGGACTTCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 221 AsnGluLeuLeuSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATAAGCTTTCAGAGAAAGGGTCAGGAGATCATTTGTCATGCTGTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGACGAGGAGTTTAAACAGAGCTTGCAGTTGGCTCATTTGGAATGAATTCAT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGCAGAAATCCTTTGATTTATGGAACACATTTCTTTAGAGGAAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluLys 320
QY 961 ACAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGGATTTGCGAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAAGCTCTTACCTTGGATGCAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 5

US-10-990-328-7547
; Sequence 7547, Application US/10990328
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USSES THEREOF
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 7547


```
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTAAATCAATGAA 480
Db 141 LeuIleAspThrTyrIleArgaspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAAGAAAGACAGATTGGCGCTTGCAGTGTAGATAGACATAGAAAA 540
Db 161 ThrMetProTyrValLysLysAlaAspThrAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGGAAGAGTGGTCCCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTGTATATTCTGGCTTAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGGCTGTATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgaspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATAAGCCTTTCAGAAAGAGGTCAGGAGATCATTTGTCATGCTGTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260
QY 781 ATTGAGCAGGAGTTTAAACAGAACCTTGCAGAGCTTGCAGTTGGCTCATTTGGAATGCAAT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACACTACATTCAGTTGCTGTAGCTGACAGATTACTTGTGAACTTCGATTCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGCAGAAAAATCCTTTTGATTTTATGAAAAACATTTCTTTAGAAAGAAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTGCAGTTATGCGAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACCTTTCACCTTGCATCAGATT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 7

```
; Sequence 94, Application US/60449629
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CANCER, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001449
; CURRENT APPLICATION NUMBER: US/60/449,629
; CURRENT FILING DATE: 2003-02-04
; NUMBER OF SEQ ID NOS: 2712
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 94
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-449-629-94
```

Alignment Scores:

Pred. No.:	1,89e-186	Length:	351
Score:	1821.00	Matches:	351
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	98.65%	Indels:	0
DB:	37	Gaps:	0

```
US-10-698-228-2 (1-1053) x US-60-449-629-94 (1-351)
QY 1 ATGGGCGACCCCGAAAGCGCGGAGCGGCTCGATCAGGATGAGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACCAACAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGGTTTGTCTTCCCAATCCAGTACCCTGATATTTGGAAATGTTAAACAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTCTGACAGCAGAGAGTGCAGTCTTATCAAGAGATCTCCCTCACTGTAACAAG 240
Db 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCATCTTACATCTTTCAGCCCTTTTTC 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTTGTAATAAATTTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTCTATGGCTTTCAAATTCATCGAGAATGTTCTCAGTGTGAGTGTACAGATTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTCAA 480
Db 141 LeuIleAspThrTyrIleArgaspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAAGAAAGACAGATTGGCGCTTGCAGTGTAGATAGAGATAAAA 540
Db 161 ThrMetProTyrValLysLysAlaAspThrAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCGCTTTCGCTGTGAGAGGAGTTTCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTGTATTTCTGGCTTAAAGAGAGGTCCTTATGCCAGGACTCACATTTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgaspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATAAGCCTTTCAGAAAGAGGTCAGGAGATCATTTGTCATGCTGTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260
QY 781 ATTGAGCAGGAGTTTAAACAGAAAGCTTGCAGAGTGGCTCATTTGGAATGAAATTCAT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACACTACATTCAGTTGCTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGCAGAAAAATCCTTTTGATTTTATGAAAAACATTTCTTTAGAAAGAAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTGCAGTTATGCGAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACCTTTCACCTTGGATGACAGATT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 8

US-60-505-218-322
 ; Sequence 322, Application US/60505218
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE REFERENCE: CL001482
 ; CURRENT APPLICATION NUMBER: US/60/505,218
 ; CURRENT FILING DATE: 2003-09-24
 ; NUMBER OF SEQ ID NOS: 22507
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 322
 ; LENGTH: 351
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-60-505-218-322

Alignment Scores:
 Pred. No.: 1,898-186 Length: 351
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.65% Indels: 0
 DB: 37 Gaps: 0

US-10-698-228-2 (1-1053) x US-60-505-218-322 (1-351)

QY 1 ATGGCGGACCCGAAAGCCGCGGCTGGATCAGATCAGAGATCATCTTCA 60
 Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
 QY 61 GACACACGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAGTCTCC 120
 Db 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerArg 40
 QY 121 CGTTTGTTCATCTTCCCAATCCAGTACCTGATATTGGAAATGTATATAACAGGCACAG 180
 Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysGlnAlaGln 60
 QY 181 GCTTCCTCTGACAGCAGAGAGGTGCGACTTATCAAGGATCTCCCTCAGTGGACACAG 240
 Db 61 AlaSerPheTrpThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 QY 241 CTTAAGCAGATGAGAGTACTTCTCTCATCTCTACATCTTAGCCCTTTTTCAGCCAGTGAT 300
 Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
 QY 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTACTCAGGAGTCCGCTTCAGAGGCT 360
 Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 QY 361 CGCTGTTTCTATGGCTTTTCAATCTCATCGAGAATGTTTCACTCAGAGATGACAGTTTG 420
 Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATCAATTGAA 480
 Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 QY 481 ACCATGCCCTATGTAAAGAAAGCAGATTGGGCTTCGAGTGGATAGCAGATAGAAAA 540
 Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 QY 541 TCTACTTTGGGAAAGAGTGGTGGCTTGTGCTGTAGAGGAGTGTCTTCTCAGGA 600
 Db 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 QY 601 TCTTTTGTCTATATTCTGGCTAAAGAGAGGTCTTATGCCAGGACTCACTTTTCC 660
 Db 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220

QY 661 AATCAACTCATCAGCAGATGAAGGACTTCTCTGTGCTTGTGCTGTGCTGTGCTTCAA 720
 Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 QY 721 TACTTAGTAATAAGCCTTCAGAAGAAAGGTGAGGAGATCATTTGATGCTGTCTCAA 780
 Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260
 QY 781 ATTGAGCAGGAGTCTTTTAAACAGAGAGCTTGCCAGTTGGCCTCATTTGGAATGAATTCATT 840
 Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 QY 841 TTGATCAAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 900
 Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 QY 901 AAGTCTTTTTCAGGAGAAATCTTTTGTGATTTTATGGAATAAATCTTTTAGAAGAAAA 960
 Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 QY 961 ACAAAATTTCTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCGAGTTTATGCGAGAAACC 1020
 Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 QY 1021 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1053
 Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 9

US-09-791-537-27671
 ; Sequence 27671, Application US/09791537
 ; GENERAL INFORMATION:
 ; APPLICANT: Bionomix, Inc.
 ; APPLICANT: Debe, Derek
 ; APPLICANT: Danzer, Joseph
 ; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
 ; FILE REFERENCE: 261/210
 ; CURRENT APPLICATION NUMBER: US/09/791,537
 ; CURRENT FILING DATE: 2001-02-22
 ; NUMBER OF SEQ ID NOS: 153055
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 27671
 ; LENGTH: 366
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-791-537-27671

Alignment Scores:

Pred. No.: 1,928-186 Length: 366
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.65% Indels: 0
 DB: 22 Gaps: 0

US-10-698-228-2 (1-1053) x US-09-791-537-27671 (1-366)

QY 1 ATGGCGGACCCGAAAGCCGCGGCTGGATCAGATCAGAGATCATCTTCA 60
 Db 16 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 35
 QY 61 GACACCAAGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAGAGTCTCGC 120
 Db 36 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 55
 QY 121 CGTTTGTCTATCTTCCATCCAGTACCTGATATTGGAAATGTATATAACAGGCACAG 180
 Db 56 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysGlnAlaGln 75
 QY 181 GCTTCCTCTGACAGCAGAGAGGTGCGACTTATCAAGGATCTCCCTCAGTGGACACAG 240
 Db 76 AlaSerPheTrpThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 95

241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTAGCCCTTTTTCAGCAGTGTAT 300
96 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaIleSerAsp 115
301 GGAATTTGTAATGAAATTTGTGGAGCCCTTTAGTCAGGAGGTCCAGAGGCT 360
116 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 135
361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGATGTTCTCATCTCAGAGATGTACAGTTTG 420
136 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 155
421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTAAATGCAATTTGAA 480
156 LeuIleAspThrTyrIleArgAspProLysLysAlaAspTyrPheLeuPheAsnAlaIleGlu 175
481 ACCATGCCCTATGTTAAGAAAAAGCAGATTCGGCCCTTCGATGGATAGCAGATAGAAAA 540
176 ThrMetProTyrValLysLysAlaAspTyrPheAlaValGluGlyValPhePheSerGly 195
541 TCTACTTTTGGGAAAGAGTGTGGCTTTCCTGCTGTAGAGAGTTCCTTCTCAGGA 600
196 SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 215
601 TCTTTTGTCTATATCTGGCTTAAAGAGAGAGTCTTATCCAGAGTCACTTTTCC 660
216 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 235
661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
236 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 255
721 TACTTAGTAATAAGCTTCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
256 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAlaValLys 275
781 ATTGAGCAGGAGTTTAAACAGAGCTTGCAGTGGCTCATTTGGAATGATTCGATT 840
276 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 295
841 TTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGAACTTCGATTCTCA 900
296 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 315
901 AAGTTTTCAGGCAGAAATCCTTTGATTTTATGAAAAAATTCCTTTTGAAGAAAA 960
316 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 335
961 ACNAATTTCTTTGAGAAAGAGTTTCAGAGTATCAGCGTTTTCAGCTTATGGCAGAAACC 1020
336 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 355
1021 ACAGATAACGCTCTTCCACCTTGGATGAGATTTT 1053
356 ThrAspAsnValPheThrLeuAspAlaAspPhe 366

RESULT 10

US-09-629-469A-12811

; Sequence 12811, Application US/09629469A

; GENERAL INFORMATION:

; APPLICANT: OTA, TOSHIO

; APPLICANT: ISOGAI, TAKAO

; APPLICANT: NISHIKAWA, TETSUO

; APPLICANT: HAYASHI, KOJI

; APPLICANT: SAITO, KAORU

; APPLICANT: YAMAMOTO, JUNICHI

; APPLICANT: ISHII, SHIZUKO

; APPLICANT: SUGIYAMA, TOMOYASU

; APPLICANT: WAKAMATSU, AI

; APPLICANT: NAGAI, KEIICHI

; APPLICANT: OTSUKI, TETSUJI

; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE

FILE REFERENCE: 084335/0123
CURRENT APPLICATION NUMBER: US/09/629, 469A
CURRENT FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: JP 1999-248036
PRIOR FILING DATE: 1999-07-29
PRIOR APPLICATION NUMBER: JP 1999-300253
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: JP 2000-118776
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: JP 2000-183767
PRIOR FILING DATE: 2000-05-02
PRIOR APPLICATION NUMBER: JP 2000-241899
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 60/159,590
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 60/183,322
PRIOR FILING DATE: 2000-02-17
NUMBER OF SEQ ID NOS: 19025
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12811
LENGTH: 351
TYPE: PRT
ORGANISM: Homo sapiens
US-09-629-469A-12811

Alignment Scores:

Pred. No.: 5,11e-186 Length: 351
Score: 1817.00 Matches: 350
Percent Similarity: 100.00% Conservatives: 1
Best Local Similarity: 99.72% Mismatches: 0
Query Match: 98.43% Indels: 0
DB: 20 Gaps: 0

US-10-698-228-2 (1-1053) x US-09-629-469A-12811 (1-351)

QY 1 ATGGGCGACCCCGAAAGCGCGGCGGTGATCAGATGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAAGAAAGTGAATTAAGTCMAATGAAGAGCACCTCTTGAAGAGAGTCTCGC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
QY 121 CGCTTTTCTCATCTTCCAAATCCAGTACCTGATATTTGGAATATTAACAGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTTCTGACACAGAGAGGTGCTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db 61 AlaSerPheTyrThrAlaGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCTCACTTTAGCCTTTTTCAGCCAGTGTAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaIleSerAsp 100
QY 301 GGAATTTGTAATGAAATTTGTGGAGCCCTTTAGTCAGGAGGTCCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCAGAAATGTTCTACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTAAATGCAATTTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCTTGCATCGATCGATAGAAAA 540
Db 161 ThrMetProTyrValLysLysAlaAspTyrPheAlaLeuArgTyrPheAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTCCTGCTGTAGAGAGTTCCTTCTCAGGA 600

Db 181 SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATCTTGGCTAAAGAGAGAGGCTTTATGTCAGAGACTCCTTTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACATCATCAGCAGAGATGAAGACTTCTACTGTGACTTTGCTTGGCTGTGATTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGCTGCTGTCTCAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260
QY 781 ATTGACGAGGAGTCTTAAACAGAGCCTTGCCAGTTGGCTCATTTGGAATGAATTCATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyValAsnCysIle 280
QY 841 TTGATGAACACATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGTTTTTCAGCAGAGAAATCTTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTTCTTTGACAAAGAGTTCAGAGTATCAGCCTTTTGCAGTTATGTCAGAGAAC 1020
Db 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTCTTCACCTTGGATGCAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 11

US-09-791-537-112185
; Sequence 112185, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonoxim, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 112185
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-112185

Alignment Scores:

Pred. No.:	5,11e-186	Length:	351
Score:	1817.00	Matches:	350
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.72%	Mismatches:	0
Query Match:	98.43%	Indels:	0
DB:	22	Gaps:	0

US-10-698-228-2 (1-1053) x US-09-791-537-112185 (1-351)

QY 1 ATGGCGCAGCCGGAAGCCGGAAGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGTTTGTCTATCTTCCATTCAGTACCTGATATTTGGAAATATTTGAACAGGCACAG 180

RESULT 12

US-10-917-503-12811
; Sequence 12811, Application US/10917503
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI

Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTTGGACAGCAGAGAGGTCGACTTATCAAGAGATCTCCCTCCTCTGGAACAG 240
Db 61 AlaserPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAAGCAGATCAGAAAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGCAGCATGAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAspSerAsp 100
QY 301 GGAATTTGTAATAATTTGGTGAGCGCTTTAGTCAGAGAGGTCAGAGTTCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGCTTCAAAATCTCATCAGAGATGTTTCTCCTCAGAGATCTACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTCGA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTTATGTTAAGAAAGAGATTTGGGCTTGGCATGGATGATGATGATGATA 540
Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGCTAGAGAGGATTTTCTTCTCAGA 600
Db 181 SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATCTTGGCTAAAGAGAGAGGCTTATGCCAGAGTCTCCTTTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACATCATCAGCAGAGATGAAGACTTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGATGCTGTCTCAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGACGAGGATTTTAAACAGAGCCTTCCAGTTCGCTCCTCATTTGGAATGAATTCATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyValAsnCysIle 280
QY 841 TTGATGAACACATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGTTTTTTCAGCAGAGAAATCTTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTTCTTTGAGAAACAGTTCAGAGTATCAGGATTCAGGATTTTTCAGAGTATG 1020
Db 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTCTTCACCTTGGATGCAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

Qy	91	GAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTTGTCTCATCTTTCCAAATCCAGTACCGCT	150
Db	69	AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis	88
Qy	151	GATATTGTGAAATATGATAACAGCACAGGCTTCCTTCGACAGCAGAAAGAGTGCAC	210
Db	89	AspIleTrpGlnMetTyrLysLysAlaGluAspPheThrAlaGluGluValAsp	108
Qy	211	TTATCAAGAGATCTCCCTCACTGGAAACAAGCTTAAGCAGATGAGAAGTACTTCACTCT	270
Db	109	LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer	128
Qy	271	CACATCTTAGCTTTTTCAGCCAGTCATGGAATTGAAATGAAATTTGGTGAGCGC	330
Db	129	HisValLeuAlaPhePheAlaSerAspGlyIleValAsnGluAsnLeuValGluArg	148
Qy	331	TTTAGTCAGGAGTGCAGGCTCCAGAGCTCGCTGTTTCTATGGCTTTCAAAATCTCATC	390
Db	149	PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet	168
Qy	391	GAGAATGTTCACACAGAGATGTACAGTTTGTGTATAGACACTTATCATCAGAGATCCCAAG	450
Db	169	GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys	188
Qy	451	AAAAGGGGAATTTTATTTAATGCAATGAAACCAATGCCTATGTTAGAAAAAAGCAGAT	510
Db	189	GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp	208
Qy	511	TGGCGCTTGGCATCGATAGCAGATAGAAAATCTACTTTTGGGAAAGAGTGGTGCCTTT	570
Db	209	TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe	228
Qy	571	GCTCCTGTAGAAGAGTTTCTTCTCAGGATCTTTTGTCTGCTATATCTTGCTAAAGAG	630
Db	229	AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys	248
Qy	631	AGAGGCTTATGCCAGGACCTCACTTTTCCAAATGCAATCATCATCAGCAGAGTGAAGACTT	690
Db	249	ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu	268
Qy	691	CACGTGCACTTTGTCCCTGATGTTCCAATACTTAGTAAATAAGCCCTTCAGAGAAAGG	750
Db	269	HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg	288
Qy	751	GTCAGGAGATCAATTGTTGATGCTGTCAAAATTCAGCAGGAGTTTTTAACAGAGCCTTG	810
Db	289	ValArgGluIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu	308
Qy	811	CCAGTTCGCCTCATTGCAATGAATTGCATTTTGTATGAACAGTACATTGAGTTGTAGCT	870
Db	309	ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla	328
Qy	871	GACAGATTACTTGTGGAACCTTGGATTTCTCAAAGCTTTTTCAGGCAGAAAAATCCTTTTCAAT	930
Db	329	AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp	348
Qy	931	TTTATGAAAAACATTTCTTTAGAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAGAG	990
Db	349	PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu	368
Qy	991	TATCAGGCGTTTTCAGTTATGTGGCAAAACACACAGATAACGCTCTTCACTTCGGATGAGAT	1050
Db	369	TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp	388
Qy	1051	TTT 1053	
Db	389	Phe 389	

RESULT 14

REC'D 14
PCT-US03-17409-171

FCI-0303-17409-171
: Sequence 171, Application PC/TUS0317409

: GENERAL INFORMATION:

APPLICANT: INCYTE CORPORATION

Qy 1051 TTT 1053
Db |||
389 Phe 389

Search completed: October 30, 2005, 06:47:55
Job time : 182.25 secs


```
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
QY 121 CGGTTTGTCATCTTCCAAATCCAGTACCTGATATTTGGAAATGATATAACAGGACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTTCGACAGCAGAGAGTGCAGTCTATCAAGAGGATCTCCCTCACCTGGAAACAG 240
Db 61 AlaserPheThrAlaGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
QY 241 CTTAAAGCAGATGAGAAGTACTCTCTCATCTTAGCCCTTTTTCGACCCAGTGTAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGTTCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTAAGAAAAAGAGATGCGGCTTCGAGTGCAGATGAGATGAGAAA 540
Db 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTyrPheAlaAlaSerArgLys 180
QY 541 TCTTACTTTTGGGAGAGAGTGGTGGCTTTGCTGTAGAGAGGATTTCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPheSerGly 200
QY 601 TCTTTTGTGCTATATCTGGCTTAAAGAGAGAGTCTTATGCCAGGATCTCACTTTTCC 660
Db 201 SerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTATCAGCAGATGAGGATCTTCACTGTGACTTGTGCTGCTGCTGCTGCTGCT 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATTAAGCTTCCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTCTCAAA 780
Db 781 ATTGACGAGAGTGTAAACAGAGCCTTGCAGATGCGCTCATTTGGAATGAATTCAT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAAACAGTACATTGAGTTGTAGCTGACAGATCTTGTGGAATCTGATTTCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGTTTTCAGGAGAGAAATCCTTTGATTTTATGAAAAAATTTCTTTAGAGAGAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTTGAGAAAGAGTTCAGAGTATCAGCGTTTGTGCGATTTATGCGAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTTCCACCTTGGATCGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 2

```
PCT-US05-21650-34
; Sequence 34, Application PC/TUS0521650
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: MIGFS AS MODIFIERS OF THE IGF PATHWAY AND METHODS OF USE
```

```
; FILE REFERENCE: EX05-017
; CURRENT APPLICATION NUMBER: PCT/US05/21650
; CURRENT FILING DATE: 2005-06-27
; PRIOR APPLICATION NUMBER: 60/581,689
; PRIOR FILING DATE: 2004-06-21
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; PCT-US05-21650-34
Alignment Scores: 51e-178 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservations: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 1 Gaps: 0
```

US-10-698-228-2 (1-1053) x PCT-US05-21650-34 (1-351)

```
QY 1 ATGGCGGACCCCGGAAAGCGCGGCGGCTGGATCAGATGAGATGATCATCTTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAAGTGAATTAAGTCAATGACAGCCACTCTCTAAGAAAGATTTCTGC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
QY 121 CGGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAATGATATAACAGGACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTTTCGACAGCAGAGAGTGCAGTCTTATCAAGAGATCTCCCTCATCTGGAACAG 240
Db 61 AlaserPheThrPheAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
QY 241 CTTAAAGCAGATGAGAAGTACTCTCTCATCTTGTAGCCTTTTTCGACCCAGTGTAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGTTCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTATCAGAGATCCCAAGAAAGGAGATTTTATTTAATGCAATTTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysAlaAspTrpAlaLeuArgTyrPheAlaAlaSerArgLys 160
QY 481 ACCATGCCCTATGTAAGAAAAAGAGATGCGGCTTCGAGTGCAGATGAGATGAGAAA 540
Db 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTyrPheAlaAlaSerArgLys 180
QY 541 TCTTACTTTTGGGAGAGAGTGGTGGCTTTGCTGTAGAGAGGATTTCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPheSerGly 200
QY 601 TCTTTTGTGCTATATCTGGCTTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
Db 201 SerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTATCAGCAGATGAGGATCTTCACTGTGACTTGTGCTGCTGCTGCTGCTGCT 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATTAAGCTTCCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTCTCAAA 780
```

```
Db      241  TyrLeuValAsnLysProSerGluGluArgValArgGluLeuLeValAspAlaValLys 260
QY      781  ATTGACGAGGAGTTTAAACAGAGCTTGCACAGTGTGGCTCATTTGGAATGAATTGCATT 840
Db      261  IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY      841  TTGATGAACACGATGAGTTGTAGCTGACAGATTACTTGTGGAACTTGGATTCTCA 900
Db      281  LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY      901  AAGGTTTTTCAGCAGAAATCCTTTGATTTTATGGAACAACTTTTGTAGAGAAA 960
Db      301  LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY      961  ACAAAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1020
Db      321  ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY      1021 ACAGATAACGTTCTTACCTTGGATGCGAGATTTT 1053
Db      341  ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 3
US-10-990-328A-7547
; Sequence 7547, Application US/10990328A
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328A
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7547
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328A-7547

Alignment Scores:
Pred. No.: 5,1e-178 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 6 Gaps: 0

US-10-698-228-2 (1-1053) x US-10-990-328A-7547 (1-351)

QY      1  ATGGCGGACCCGGAAGCCGAGCGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Db      1  MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY      61  GACACCAACGAGTGAATTAAGTCAATGAAGCCACTCTTAAGAAGAGTTCTCC 120
Db      21  AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
QY      121  CGGTTTGTTCATCTTCCATCCAGTACCTGATATTGTGAAAATGTATAACAGGCACAG 180
Db      41  ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY      181  GCTTCCTTTCGACAGCAGAGAGGTTCGACTTATCAAGGATCTCCCTCACTGGACACAG 240
Db      61  AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
QY      241  CTTAAAGCAGATGACAGAGTACTTCTCATCTCATCTTACCTTTTTCGACGCCAGTGT 300
Db      81  LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY      301  GGAATTTGAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCCAGAGGCT 360
Db      301  GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY      361  CGCTGTTTCTATGCTTCAAATCTCATCAGAGATGTTCACTCAGAGATGTACAGTTG 420
Db      121  ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY      421  CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAAATCAATGAA 480
Db      141  LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY      481  ACCATGCCCTATGTAAGAAAAAGCAGATTGGGCTTGGCGATGATGATGATGATGATA 540
Db      161  ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY      541  TCTACTTTTGGGAAAGAGTGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
Db      181  SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY      601  TCTTTTCTGCTATATTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTCTCC 660
Db      201  SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY      661  AATCAACTCATCAGCAGAGATGAAGACTTCTCACTGTGACTTTGCTTGCCTGATGTTCCA 720
Db      221  AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY      721  TACTTAGTAATAAGCTTCTCAGAAAGAGGTGAGGAGATCATTTGATGATGATGATGAT 780
Db      241  TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY      781  ATTGAGCAGGAGTTTAAACAGAGCTTGGCCAGTGGCCCTCATTTGGAATGAATTGCATT 840
Db      261  IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY      841  TTGATGAACACGATGATGAGTTTGTAGCTGACAGATTTACTTGTGGAATCTTGCATTCTCA 900
Db      281  LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY      901  AAGTTTTTTCAGCAGAAATCCTTTGATTTTATGGAACAACTTTTGTAGAGAAA 960
Db      301  LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY      961  ACAAAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1020
Db      321  ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY      1021 ACAGATAACGCTTCTTACCTTGGATGCGAGATTTT 1053
Db      341  ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 4
US-11-088-686-15
; Sequence 15, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-15

Alignment Scores:
Pred. No.: 5,1e-178 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
```



```
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 7 Gaps: 0

US-10-698-228-2 (1-1053) x US-11-088-686-15 (1-351)
QY 1 ATGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGAGATCATCTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAAGTAAATGAATCAATGAAGAGCCACTCTCTAAGAAAGTCTCGC 120
DB 21 AspThrAsnGluSerGluLeuIleYsSerAsnGluProLeuLeuArgYsSerArg 40
QY 121 CGGTTTGTCACTTTCCATCCAGTACCTGATATTTGGAAATCTATAAAGAGCAGAG 180
DB 41 ArgPheValIlePheProGlnGlyProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTCTGACAGCAGAGAGGTCCGACTTATCAAGGATCTCCCTCAGTGGAAACAG 240
DB 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTAAAGCAGATGAAAGTCACTCTCACTCATCTCTCACTCTCTCTCTCTCTCTCTCT 300
DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAlaSerAsp 100
QY 301 CGAATTGTAATGAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnProGluAla 120
QY 361 CGCTGTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
DB 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATPAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATCAATTTGAA 480
DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTAAGAAAGAGGATTTGGCCCTTGGATGATGATGATGATGATGAT 540
DB 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTGGGGAAGAGTGGCTTTGCTGCTGTAGAAGAGTCTTCTCTCTCAGCA 600
DB 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPheSerGly 200
QY 601 TCTTTTGTGCTATATCTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC 660
DB 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTCATCAGCAGATGAGGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 720
DB 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATAAGCTTCAAGAAAGGCTCAGGAGATCATTTGTCATGCTGTCAAA 780
DB 241 TyrLeuValAsnLysProSerGluArgValArgGluIleValAspAlaValLys 260
QY 781 ATTAGCAGGAGGTTTTTAAAGAGGCTTGGCCCTCATTTGGAATGAATTTGCAATT 840
DB 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
DB 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCCGCGAGAAATCTTTTGAATTTATCGAAACATTTCTTTAGAGGAAA 960
DB 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTTGAGAAACAGTTTTTCAGAGTATCAGCGTTTTGCACTTTATGGCAGAAACC 1020
DB 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
```

```
QY 1021 ACAGATAACGCTCTTCCACTTGGATGAGATTTT 1053
DB 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 5

```
PCT-US05-09639-17
; Sequence 17, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004WOI
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-17
```

```
Alignment Scores:
Pred. No.: 1,31e-177 Length: 351
Score: 1817.00 Matches: 350
Percent Similarity: 100.00% Conservatives: 1
Best Local Similarity: 99.72% Mismatches: 0
Query Match: 98.43% Indels: 0
DB: 1 Gaps: 0
```

US-10-698-228-2 (1-1053) x PCT-US05-09639-17 (1-351)

```
QY 1 ATGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGAGATCATCTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAAGTCAATGAATGAAGAGCCACTCTCTAAGAAAGTCTCGC 120
DB 21 AspThrAsnGluSerGluLeuIleYsSerAsnGluProLeuLeuArgYsSerSerArg 40
QY 121 CGGTTTGTCTATCTTTCCAAATCCAGTACCTGATATTTGGAAATCTATAAAGAGCAGAG 180
DB 41 ArgPheValIlePheProGlnGlyProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTCTGACAGCAGAGAGGTTCGATTTCAAGGATCTCCCTCAGTGGAAACAG 240
DB 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTAAAGCAGATGAAAGTCACTCTCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAlaSerAsp 100
QY 301 CGAATTGTAATGAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnProGluAla 120
QY 361 CGCTGTTCTATGGCTTTCAAAATTTCTCATCCAGAAATTTCTCTCAGAGATGTACAGTTTG 420
DB 121 ArgCysPhePheGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATPAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 480
DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTAAGAAAGAGGATTTGGCCCTTGGATGATGATGATGATGATGAT 540
DB 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTGGGGAAGAGTGGCTTTGCTGCTGTAGAAGAGTCTTCTCTCTCAGCA 600
DB 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPheSerGly 200
```

QY 601 TCTTTTGGCTGCTATATCTTGGCTAAAGAGAGAGAGTCTTTATGCGAGAGTCACTTTTCC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACATCATCAGCGAGATGAGAGCTTCACTGTGAGCTTTGCTTGGCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATCTTGTATGCTGTCMAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGAGCAGAGTCTTTTAAACAGAGCCTTGCAGTGGCTCATTTGGAATGAATTCATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACACATGACATGAGTTTGTGCTGACAGATTAATCTTGTGAACTTGGATTCCTCA 900
Db 281 LeuMetLysGlnTyrlleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGCTTTTTCAGCAGAGAAATCCTTTTGTATTTATGGAATAATCTTTTGAAGAAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTCTTTGAGAAACGAGTTTCAGAGATATCAGCGTTTTCAGTGTATGCGAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrlleGluArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 6

US-11-088-686-17

; Sequence 17, Application US/11088686

; GENERAL INFORMATION:

; APPLICANT: Yen, Yun

; TITLE OF INVENTION: DRUG SCREENING

; FILE REFERENCE: 14037-004001

; CURRENT APPLICATION NUMBER: US/11/088,686

; CURRENT FILING DATE: 2005-03-23

; PRIOR APPLICATION NUMBER: US 60/556,836

; PRIOR FILING DATE: 2004-03-25

; NUMBER OF SEQ ID NOS: 59

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 17

; LENGTH: 351

; TYPE: PRT

; ORGANISM: Homo sapiens

US-11-088-686-17

Alignment Scores:

Pred. No.:	1,316-177	Length:	351
Score:	1817.00	Matches:	350
Percent Similarity:	100.00%	Conservative:	1
Best local Similarity:	99.72%	Mismatches:	0
Query Match:	98.43%	Indels:	0
DB:	7	Gaps:	0

US-10-698-228-2 (1-1053) x US-11-088-686-17 (1-351)

QY 1 ATGGGCGACCCGGAAGCGCGGAGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAAGCAAGTGAATAAGTCAATGAGAGCCACTCTTAAGAAAGAGTTCGC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerArg 40
QY 121 CGGTTTGTTCATCTTTCCATCAGTACCTGCTGATATTTGGAAATGATATAACAGGACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrlleProAspIleTrpLysGlnAlaGln 60

QY 181 GCTTCTCTTCGAGCAGCAGAGAGTGCAGCTTATCAAAAGGATCTCCCTCACTGGAACAAG 240
Db 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAAGCAGATGAGAGTACTTCACTCTCAATCTTACATCTTCTAGCCTTTTTCGACCCAGTGA 300
Db 81 LeuLysAlaAspGluLysTyrlleSerHisIleLeuAlaPhePheAlaAspAsp 100
QY 301 GGAATTCTAATGAAATTTGGTGGAGCGCTTTTGTAGTCCAGGAGTGCAGGTTCCAGAGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGCTCTTCAAAATCTCATCGAGAAATGTTCTACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPhePheGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrlleSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTAAATGCAATTGAA 480
Db 141 LeuIleAspThrTyrlleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATGGCGCTTGCATGATGATGATGATGATGATGAT 540
Db 161 ThrMetProTyrlleValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGCTGCTAGAGAGGATTTTCTTCTCAGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTCTATATTCTGGCTAAAGAGAGAGTCTTATGCCAGGAGTCACTTTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACATCATCAGCAGATGAGAGCTTCACTGTGACTTTTGTGCTGCTGCTGCTGCTGCT 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATCTTGTGCTGCTGCTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGAGCAGGAGTCTTTTAAACAGAGCCTTGCAGTGGCTTCCCTCATTTGGAATGAATTCATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACACAGTACATGAGTTTGTAGCTGACAGATTAATCTTGTGGAACCTTGGATTCCTCA 900
Db 281 LeuMetLysGlnTyrlleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGCTTTTTCAGCAGAGAAATCCTTTTGTATTTTATGGAATAATCTTTTGAAGAAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTGTATGCGAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrlleGluArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 7

PCT-US05-09639-19

; Sequence 19, Application PC/TUS0509639

; GENERAL INFORMATION:

; APPLICANT: Yen, Yun

; TITLE OF INVENTION: DRUG SCREENING

; FILE REFERENCE: 14037-004001

; CURRENT APPLICATION NUMBER: PCT/US05/09639

; CURRENT FILING DATE: 2005-03-23

; PRIOR APPLICATION NUMBER: US 60/556,836

; PRIOR FILING DATE: 2004-03-25

; NUMBER OF SEQ ID NOS: 59

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-19

Alignment Scores:
Pred. No.: 1.66e-177
Score: 1846.00
Percent Similarity: 100.00%
Best Local Similarity: 99.72%
Query Match: 98.37%
DB: 1
Length: 351
Matches: 350
Conservative: 1
Mismatch: 0
Indels: 0
Gaps: 0

```

US-10-698-228-2 (1-1053) x PCT-US05-09639-19 (1-351)

Qy	1	ATGGGGACCCGGAAAGCCGGAAGCGCGCGGTGGATCAGGATCAGAGATCATCTTCA	60
Db	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer	20
Qy	61	GACACCAACGAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGATTCTCGC	120
Db	21	AspThrAsnGluSerGluLeuLysSerAsnGluProLeuLeuArgLysSerSerArg	40
Qy	121	CGGTTTCTCATCTTTCCAATCCAGTACCCCTGATATTTCGAAATGTATATAACAGGCACAG	180
Db	41	ArgPheValIlePheProIleGlnTyProAspIleTrpLysMetTyLysGlnAlaGln	60
Qy	181	GCTTCTCTTGGACAGCAGAAAGAGGTGCACTTATCAAAAGGATCTCCCTCAGTCGGAACAAG	240
Db	61	AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys	80
Qy	241	CTTAAAGCAGATGAGAAGTACTTCATCTCTCACATCTTAGCCCTTTTTCGACCCAGTGTAT	300
Db	81	LeuLysAlaAspGluLysTyPheLleSerHisIleLeuAlaPhePheAlaAlaSerAsp	100
Qy	301	GGAATTCTAAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGTGCGAGGTCCTCAGAGGCT	360
Db	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla	120
Qy	361	CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGTGTACAGTTTG	420
Db	121	ArgCysPheTrpGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTySerLeu	140
Qy	421	CTGATACACACTTACATCAGAGATCCCAAGAAAGGGAAATTTTATTAATGCAATTGAA	480
Db	141	LeuIleAspThrTyrlleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160
Qy	481	ACCATGCCCTATGTTAGAAAAAGCAGATTGGGCTTTCGGATGGATAGCAGATGAAAA	540
Db	161	ThrMetProTyValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys	180
Qy	541	TCTACTTTTGGGAAAGAGTGGTGGCCCTTGCTGCTGTAGAGGAGTTTCTTCTCAGCA	600
Db	181	SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly	200
Qy	601	TCTTTTCTGCTATATTCTGGCTAAAGAAAGAGGTCTTATGCCAGGACTCACTTTTTC	660
Db	201	SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer	220
Qy	661	AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA	720
Db	221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln	240
Qy	721	TACTTTAGTAAATAAGCCCTTCAGAGAAGAGGTTCAGCGAGATCATTTGTCGTGTCAAA	780
Db	241	TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys	260
Qy	781	ATTGAGCAGGAGTTTAAACAGAGCCCTTGGCAGTTGGCCTCATTTGGAATGATTCATT	840
Db	261	IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle	280

```

Qy 841 TTGATGAACAGTACATTGAGTTTCTAGCTGCAGAGATTACTTGTGGAACTTGGATTCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Qy 901 AAGGTTTTTTCAGCGAGAAAATCCTTTTGGATTTTTATGGAACAACTTTCTTTAGAGGAAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Qy 961 ACAAAATTTCTTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGCTTATGGCAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Qy 1021 ACAGATAACGCTCTTCACCTTGGATGCAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 8
US-11-088-686-19
; Sequence 19, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-19

Alignment Scores:
Pred. No.: 1,66e-177 Length: 351
Score: 1816.00 Matches: 350
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.72% Mismatches: 0
Query Match: 98.37% Indels: 0
DB: 7 Gaps: 0

US-10-698-228-2 (1-1053) x US-11-088-686-19 (1-351)

```

QY	1	ATGGGGACCCGGAAGCGCGCGCTCGATCAGATGAGAGATCATTTCA	60
DB	1	MetGlyAspProGluuArgProGluuAlaIaGlyLeuAspGlnAspGluuArgSerSer	20
QY	61	GACACCAACAAAGTCGAATAAAGTCNAATGAGAGCCCTCTCTAGAAAGAGTCTTCGC	120
DB	21	AspThrAsnGluSerGluIleIysSerAsnGluGluProleuIeuuArgIysSerSerArg	40
QY	121	CGGTTTCTCATCTTTCCAATCCCAAGTACCTCGATATTGGAAAAATGTATAACAGGCACAG	180
DB	41	ArgPheValIlePheProIleGlnItyrProAspIleItyrIysMetItyrIysGlnAlaGln	60
QY	181	GCTTCTCTTCGACAGCAGAAGAGGTGCGACTTATCAAAAGGATCTCCCTCACTCGAACCAAG	240
DB	61	AlaSerPheItyrThrAlaGluGluValAspLeuSerIysAspLeuProHisItyrAsnIys	80
QY	241	CTTAAACGACATGAGAAGTACTTCATCTCAGATCTTAGCCCTTTTTCGAGCCAGCTGAT	300
DB	81	LeuIysAlaAspGluIlystyrPheIleSerHisIleLeuAlaPhePheAlaIaSerAsp	100
QY	301	GGAAATCTAAATCAAAATTTGGTGGACGCTTTAGTCAGAGGTGCGAGGTTCACAGAGCT	360
DB	101	GlyIleValAsnGluAsnLeuValGluuArgPheSerGlnGluValGlnValProGluAla	120
QY	361	CGCTGTTTCTATGGCTTTTCAAAATCTCATTCGAAATGTTTCACTCAGAGATGTACAGTTTG	420
DB	121	ArgCysPheItyrGlyPheGlnIleLeuIleLeuAsnValHisSerGluMetItyrSerLeu	140

```
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATGAA 480
Db 141 LeuileAspThrTyrileArgAspProLysLysArgGluPheLeuPheAsnAlaileGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTGGCCCTTCGGATCGATACAGATAGAAA 540
Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpileAlaAspArgLys 180
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTGTCTGTAGTAGAGGAGTCTTCTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValaGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATCTTGGCTAAAGAGAGAGTCTTATGCCAGACTCCTTTTCTCC 660
Db 201 SerPheAlaAlailePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTCATCAGCAGATGAGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 720
Db 221 AsnGluLeuileSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAGCTTTCAGAGAAAGGTCAGGAGATCATTTGTCATGCTGTCAA 780
Db 241 TyrLeuValAsnLysProSerGluArgValaArgGluileValaAspAlaValLys 260
QY 781 ATTGAGCAGAGTCTTAAAGAGCTTTCAGAGTTCAGTGTGACTTTTGTCTGATGCTCAA 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuileGlyMetAsnCysIle 280
QY 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTTCTCA 900
Db 281 LeuMetLysGlnTyrileGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGCAGAAATCCTTTTGTATTTATGGAACACATTTCTTTAGAGGAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnLysSerLeuGluGlyLys 320
QY 1021 ACAGATAACGTCCTTACCTTGGATGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 9

```
PCT-US05-09639-21
; Sequence 21, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen. Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-21
```

```
Alignment Scores:
Pred. No.: 3,38e-177 Length: 351
Score: 1813.00 Matches: 349
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.43% Mismatches: 0
Query Match: 98.21% Indels: 0
DB: 1 Gaps: 0
```

US-10-698-228-2 (1-1053) x PCT-US05-09639-21 (1-351)

```
QY 1 ATGGCGCACCAGAAAGCCGCGGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAaaglyLeuAspGlnAspGluArgSerSerSer 20
QY 61 GACACCAACGAAAGTGAATTAAGTCAATTAAGAGACCACTCCTTAAGAAAGATCTCGC 120
Db 21 AspThrAsnGluSerGluileLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
QY 121 CGGTTTGTCTATCTTCCAAATCCAGTACCTCGATATTTGGAAATGTATAACACGCACAG 180
Db 41 ArgPheValilePheProileGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTTCTGGACAGCAGAGAGGTCGACTTATCAAAAGGATCTCCCTCATCTGGAACAAG 240
Db 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAAGCAGATCAGAGTACTTCTCATCTCTCATCATCTTACGCTTTTGTGACGAGTAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisileLeuAlaPhePheAlaAspAsp 100
QY 301 GGAATTTTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATCGCTTTCAAATTTCTCATCGAATGTTCTCATCAGAGATGTACAGTTG 420
Db 121 ArgCysPhePheGlyPheGlnIleLeuileGluAsnValHisSerGluMetPheSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATGAA 480
Db 141 LeuileAspThrTyrileArgAspProLysLysArgGluPheLeuPheAsnAlaileGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTGGCCCTTCGGATCGATACAGATAGAAA 540
Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpileAlaAspArgLys 180
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTGTCTGTAGAGGATTTCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValaAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGGCTTATGCGCAGGACTCACATTTTTC 660
Db 201 SerPheAlaAlailePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTCATCAGCAGATGAGACTTCACTGTGACTTTTGTCTGCTGATTTCCAA 720
Db 221 AsnGluLeuileSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAGCTTTCAGAGAAAGGTCAGGAGATCATTTGTCATGCTGTCAA 780
Db 241 TyrLeuValAsnLysProSerGluArgValaArgGluileValaAspAlaValLys 260
QY 781 ATTGAGCAGAGTCTTAAAGAGCTTTCAGAGTTCAGTGTGACTTTTGTGAACTGATTT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuileGlyMetAsnCysIle 280
QY 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTTCTCA 900
Db 281 LeuMetLysGlnTyrileGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGCAGAAATCCTTTTGTATTTATGGAACACATTTCTTTAGAGGAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnLysSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTGGAGAAACAGTTTCAGAGTATCAGCGTTTGTGAGTATGCGAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGTCCTTACCTTGGATGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 10

```
US-11-088-686-21
; Sequence 21, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-21

Alignment Scores:
Pred. No.: 3,38e-177 Length: 351
Score: 1813.00 Matches: 349
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.43% Mismatches: 0
Query Match: 98.21% Indels: 0
DB: 7 Gaps: 0

US-10-698-228-2 (1-1053) x US-11-088-686-21 (1-351)

QY 1 ATGGGCGACCGAAGCGCGGAGCGGCTGGATCAGATCAGATCATCTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAGTCAATTAAGTCAATGAAGCCCTCTCAAGAAAGTCTCGC 120
DB 21 AspThrAsnGluSerGluLeuValAspLeuSerLysAspLeuProHisTrpAsnLys 40
QY 121 CGTTTGTTCATCTTCCATCCAGTACCTCTGATATTGGAATGTAATAACAGCACAG 180
DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTCGACAGAGAGGTGCGATTCATCAAGAGTCTCCCTCACTCGAACAG 240
DB 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAGACGATGAGAGTACTCTCATCTCTACATCTTACCTCTTGTGAGCCAGCTGAT 300
DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
QY 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTCTATGGCTTTCAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 420
DB 121 ArgCysPhePheGlyPheGlnIleLeuIleGluAsnValHisSerGluMetPheSerLeu 140
QY 421 CTGATACACATTACATCAGAGATCCCAAGAAAGGAAATTTTATTATGCAATGAA 480
DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTGGCGCTTGGATGGATAGCAGATAGAAA 540
DB 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGCTTGTGCTGTAGAGAGGATTTTCTTCTCAGGA 600
DB 181 SerThrPheGlyGluArgValAlaPheAlaValGluGluValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCCTTTTCC 660
DB 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACATCATCAGCAGATGAAGACTTTCACCTGTGACTTTTCTGCTGATGTTCCAA 720

RESULT 11
PCT-US05-09639-23
; Sequence 23, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-23

Alignment Scores:
Pred. No.: 4,29e-177 Length: 351
Score: 1812.00 Matches: 349
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.43% Mismatches: 0
Query Match: 98.16% Indels: 0
DB: 1 Gaps: 0

US-10-698-228-2 (1-1053) x PCT-US05-09639-23 (1-351)

QY 1 ATGGGCGACCGAAGCGCGGAGCGGCTGGATCAGATCAGATCATCTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAGTCAATTAAGTCAATGAAGCCCTCTCAAGAAAGTCTCGC 120
DB 21 AspThrAsnGluSerGluLeuValAspLeuSerLysAspLeuProHisTrpAsnLys 40
QY 121 CGTTTGTTCATCTTCCATCCAGTACCTCTGATATTGGAATGTAATAACAGCACAG 180
DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTCGACAGAGGTGCGATTCATCAAGAGTCTCCCTCACTCGAACAG 240
DB 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAGACGATGAGAGTACTCTCATCTCTACATCTTACCTCTTGTGAGCCAGCTGAT 300
DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
QY 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTCTATGGCTTTCAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 420
DB 121 ArgCysPhePheGlyPheGlnIleLeuIleGluAsnValHisSerGluMetPheSerLeu 140
QY 421 CTGATACACATTACATCAGAGATCCCAAGAAAGGAAATTTTATTATGCAATGAA 480
DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTGGCGCTTGGATGGATAGCAGATAGAAA 540
DB 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGCTTGTGCTGTAGAGAGGATTTTCTTCTCAGGA 600
DB 181 SerThrPheGlyGluArgValAlaPheAlaValGluGluValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCCTTTTCC 660
DB 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACATCATCAGCAGATGAAGACTTTCACCTGTGACTTTTCTGCTGATGTTCCAA 720

221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
721 TACTTAGTAATAAGCTTCAGAGAAGAGGTGAGGAGATCATTTGTTGCTGTCAA 780
241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
781 ATTGAGCAGGAGTGTAAACAGAGCCTTCCAGTTCGCTCATTTGGAATCAATTCATT 840
261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
841 TTGATGAAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
901 AAGGTTTTCAGGCAGAAATCTTTTGTGTTTATGTAATGAAACATTTCTTTAGAGGAAA 960
301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
961 ACAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGATTATGCCAGAAACC 1020
321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
1021 ACAGATAACGCTCTTCACCTTGGATGCGAGATTTT 1053
341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

```
Db      81  LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaLaserAsp 100
QY      301  CGAATTGTAATAAATTTGGTGGAGCGCTTTAGTCAGGAGTTCAGAGGCT 360
Db      101  GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValProGluAla 120
QY      361  CGCTGTTCTATGGCTTCCAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 420
Db      121  ArgCysPhePheGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTrpSerLeu 140
QY      421  CTGATAGACATTCATCAGATCCCAAGAAAGGGAATTTTATTTATGCAATGAA 480
Db      141  LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY      481  ACCATGCCCTATGTTAAGAAAAGCAGATTGGCGCTTCGATGAGTACAGATAGAAA 540
Db      161  ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY      541  TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAAGAGTTCCTTCTCAGGA 600
Db      181  SerThrPheGlyGluArgValValAlaPheAlaIleValGluGlyValPhePheSerGly 200
QY      601  TCTTTTGTGCTATATTTCTGGCTTAAGAAGAGAGTCTTTATGCCAGACTCACTTTTCC 660
Db      201  SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY      661  AATGAACATCATCAGCAGATGAGGACTTCCTACTGTGACTTTGCTTGCCTGATGTCCAA 720
Db      221  AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY      721  TACTTAGTAAATAAGCTTCAGAGAAGAGGTCAGGAGATCATCTTCTGCTGTCAAA 780
Db      241  TyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAlaValLys 260
QY      781  ATTGACGAGGAGTTCCTTAAAGAGAGCTTGCAGTTCCTTATGCAATGCAATTCAT 840
Db      261  IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY      841  TTGATGAACAGTACATTTAGTTGTAGCTGACAGATTACTTTGTGAACCTTGGATTCTCA 900
Db      281  LeuMetLysGlnTyrIleGluPheValAlaAspPheMetGluAsnIleSerLeuGluLys 300
QY      901  ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGTCAGAAACC 1020
Db      321  ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY      1021  ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1053
Db      341  ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 12

```
PCT-US05-09639-25
; Sequence 25, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-25
```

```
Alignment Scores:
Pred. No.: 4,29e-177 Length: 351
Score: 1812.00 Matches: 349
Percent Similarity: 100.00% Conservations: 2
Best Local Similarity: 99.43% Mismatches: 0
Query Match: 98.16% Indels: 0
DB: 1 Gaps: 0
US-10-698-228-2 (1-1053) x PCT-US05-09639-25 (1-351)
```

```
QY      1  ATGGGCGACCCGAAAGCCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db      1  MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY      61  GACACCAACGAAAGTGAATAAGTCAATCAAGAGCCACTCCTAAGAAAGTTCTGCG 120
Db      21  AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
QY      121  CGGTTTGTCTATCTTCCCAATCCAGTACCCTCATATTTGGAAAAATGTATAACAGGCACAG 180
Db      41  ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY      181  GCTTCTCTTCTGAGCAGCAGAGAGTCCGACTTATCAAGAGGATCTCCCTCACTGGGAACAG 240
Db      61  AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY      241  CTTAAAGCAGATGAGAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCGACCCAGTGAT 300
Db      81  LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAspSerAsp 100
QY      301  GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGTTCAGAGGCT 360
Db      101  GlyIleValAsnGluAsnLeuValArgPheSerGlnGluValGlnValProGluAla 120
QY      361  CGCTGTTTCTATGCTTCAAAATTTCTCATCAGAAATGTTCTACTCAGAGATGTACAGTTGTG 420
Db      121  ArgCysPheTrpGlyPheGlnIleLeuIleGluAsnValHisSerGluMetPheSerLeu 140
QY      421  CTGATAGACATTCATCAGAGATCCCAAGAAAGGGAATTTTATTTATGCAATGAA 480
Db      141  LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY      481  ACCATGCCCTATGTTAAGAAAAGCAGATTGGCGCTTCGATGAGTACAGATAGAAA 540
Db      161  ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY      541  TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAAGAGTTCCTTCTCAGGA 600
Db      181  SerThrPheGlyGluArgValValAlaPheAlaIleValGluGlyValPhePheSerGly 200
QY      601  TCTTTTGTGCTATATTTCTGGCTTAAGAAGAGAGTTCCTTATGCCAGGACTCACTTTTTC 660
Db      201  SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY      661  AATGAACATCATCAGCAGATGAGGACTTCCTACTGTGACTTTGCTTGCCTGATGTCCAA 720
Db      221  AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY      721  TACTTAGTAAATAAGCTTCAGAGAAGAGGTCAGGAGATCATCTTGTGCTGTCAAA 780
Db      241  TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY      781  ATTGACGAGGAGTTCCTTAAAGAGAGCTTGCAGTTCCTTATGCAATGCAATTCAT 840
Db      261  IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY      841  TTGATGAACAGTACATTTAGTTGTAGCTGACAGATTACTTTGTGAACCTTGGATTCTCA 900
Db      281  LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY      901  AAGGTTTTCAGGCAGAAATCTTTTGTGATTTTATGGAACATTTCTTTTGAAGGAAA 960
```



```
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyIlys 320
QY 961 ACAATTTCTTGGAGAACGAGTTTCAGAGTATCAGCGTTTGGCTGCTAGAGGAGTTTCTTCTCAGGA 600
Db 321 ThrAsnPheGluGlyArgValSerGluGlyGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGTTTCCACCTTGGATCGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 13
US-11-088-686-23
; Sequence 23, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-23

Alignment Scores:
Pred. No.: 4,29e-177 Length: 351
Score: 1812.00 Matches: 349
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.43% Mismatches: 0
Query Match: 98.16% Indels: 0
DB: 7 Gaps: 0

US-10-698-228-2 (1-1053) x US-11-088-686-23 (1-351)
QY 1 ATGGCGACCCGGAAGCGGAGCGGCTGGATCAGATGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAAGGAAGTAAAGTCAATGAAGACCACTCTTAAGAAAGAGTTCTCGC 120
Db 21 AspThrAsnGluSerGluIleYsSerAsnGluGluProLeuLeuArgIysSerArg 40
QY 121 CGTTTGTTCATCTTCCAAATCCAGTACCTGTATTTGGAAAATGTATAACAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyProAspIleTrpLysMetTyIysGlnAlaGln 60
QY 181 GCTTCTCTTGACAGCAGACAGAGTGCAGTCTATCAAGGATCTCCCTCAGTGGAAACAG 240
Db 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAGCAGATGAGAAGTACTTCTCTCATCTTAGCTTTTTCAGCCCTTTTTCAGCCAGTGT 300
Db 81 LeuLysAlaAspGluLysTyPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
QY 301 GGAATGTAAATGAAAATTTGGTGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTTTCTATGGCTTTCAAAATCTCATCGAGATGTTTCATCAGAGATGATCAGTTTG 420
Db 121 ArgCysPhePheGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTrpSerLeu 140
QY 421 CTGATGACACTTACATCAGATCCCAGAAAGGAATTTTATTTAATGCAATTGAA 480
Db 141 LeuIleAspThrTyIleArgAspProLysIysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTTGTAAAGAAAAAGCAGATTTGGGCTTCCGATGGATAGCAGATAAAAA 540
Db 161 ThrMetProTyTrValLysLysIysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAGAAAGAGTGGCTTTGCTGCTGCTAGAGGAGTTTCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACATTTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAATCATCAGCAGAGATGAAGACTTCACCTGCTGCTGCTTGGCTTGCCTGATTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGTTCAGGAGATCATTTGTGTGCTGCTCAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGACGAGGAGTTTAAACAGAACGCTTGCAGTTCGCTCATTTGCAATGCAATTCATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAAACAGTACATTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 900
Db 281 LeuMetLysGlnTyIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGGAACACATTTCTTTAGAAGGAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTCTTTCAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTCAGTTCAGTTCAGTTC 1020
Db 321 ThrAsnPheGluGlyLysArgValSerGluTyGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGTTTCCACCTTGGATCGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 14
US-11-088-686-25
; Sequence 25, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-25

Alignment Scores:
Pred. No.: 4,29e-177 Length: 351
Score: 1812.00 Matches: 349
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.43% Mismatches: 0
Query Match: 98.16% Indels: 0
DB: 7 Gaps: 0

US-10-698-228-2 (1-1053) x US-11-088-686-25 (1-351)
QY 1 ATGGCGACCCGGAAGCGGAGCGGCTGGATCAGATGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAAGGAAGTAAAGTCAATGAAGACCACTCTTAAGAAAGAGTTCTCGC 120
Db 21 AspThrAsnGluSerGluIleYsSerAsnGluGluProLeuLeuArgIysSerArg 40
QY 121 CGTTTGTTCATCTTCCAAATCCAGTACCTGTATTTGGAAAATGTATAACAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyProAspIleTrpLysMetTyIysGlnAlaGln 60
QY 181 GCTTCTCTTGACAGCAGACAGAGTGCAGTCTATCAAGGATCTCCCTCAGTGGAAACAG 240
Db 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAGCAGATGAGAAGTACTTCTCTCATCTTAGCTTTTTCAGCCCTTTTTCAGCCAGTGT 300
Db 81 LeuLysAlaAspGluLysTyPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
QY 301 GGAATGTAAATGAAAATTTGGTGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTTTCTATGGCTTTCAAAATCTCATCGAGATGTTTCATCAGAGATGATCAGTTTG 420
Db 121 ArgCysPhePheGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTrpSerLeu 140
QY 421 CTGATGACACTTACATCAGATCCCAGAAAGGAATTTTATTTAATGCAATTGAA 480
Db 141 LeuIleAspThrTyIleArgAspProLysIysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTTGTAAAGAAAAAGCAGATTTGGGCTTCCGATGGATAGCAGATAAAAA 540
Db 161 ThrMetProTyTrValLysLysIysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAGAAAGAGTGGCTTTGCTGCTGCTAGAGGAGTTTCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACATTTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAATCATCAGCAGAGATGAAGACTTCACCTGCTGCTGCTTGGCTTGCCTGATTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGTTCAGGAGATCATTTGTGTGCTGCTCAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGACGAGGAGTTTAAACAGAACGCTTGCAGTTCGCTCATTTGCAATGCAATTCATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAAACAGTACATTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 900
Db 281 LeuMetLysGlnTyIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGGAACACATTTCTTTAGAAGGAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTCTTTCAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTCAGTTCAGTTCAGTTC 1020
Db 321 ThrAsnPheGluGlyLysArgValSerGluTyGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGTTTCCACCTTGGATCGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

```
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
QY 121 CGGTTTGTTCATCTTCCAAATCCAGTACCCTGATATTTGGAAATGTATAAAGCAGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTCTGACAGCAGAGAGGTGCACTTATCAAAGGATCTCCCTCAGTGGAAACAAG 240
Db 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrPheAsnLys 80
QY 241 CTTAAGCAGATGAGAAGTACTCTCATCTCATCTTACCTTTCAGGCTTTTTCAGCCAGTGAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 CGAATTTGTAATGAAATTTGGTGAGGCTTTAGTCAGAGGTGCGAGTTCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTCAAATCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetPheSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTATGCAATTTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTTAAAGAAAAGCAGATTGGGCTTCGAGTGTAGTACAGATAGAAA 540
Db 161 ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAGAGAGGATTTTCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATCTGGCTTAAAGAGAGAGTCTTATCCAGGACTCACTTTTTC 660
Db 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACTCATCAGCAGATGAGGACTTCACTGTGCTTTCAGTGTGCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGCTCAGGAGATCATTTGTGTGTCTCAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260
QY 781 ATGAGCAGGAGTTTAAACAGAACCTTGCAGTTCGCTCATTTGGAATGAATTCATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAAACAGTACATTTGAGTTGTAGCTGACAGATTCATTTGTGGAACCTTGATTTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGGAGAGAAATCTTTTCAATTTATGGAACATTTCTTTAAGAGGAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTCTTTGAAACAGGATTTTCAGAGTATCAGCGTTTTCAGTTTATGCGAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTTTCACCTTGATGTCAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 15

PCT-US05-09639-27

```
; Sequence 27, Application pc/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004WO1
```

```
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-27
```

Alignment Scores:

Pred. No.:	5,43e-177	Length:	351
Score:	1811.00	Matches:	349
Percent Similarity:	100.00%	Conservative:	2
Best Local Similarity:	99.43%	Mismatches:	0
Query Match:	98.10%	Indels:	0
DB:	1	Gaps:	0

US-10-698-228-2 (1-1053) x PCT-US05-09639-27 (1-351)

QY	1	ATGGGCGACCCGAAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTTCA	60
Db	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer	20
QY	61	GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCCTAAGAAAGAGTCTCGC	120
Db	21	AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSer	40
QY	121	CGGTTTGTCTATCTTCCAAATCCAGTACCCTGATATTTGGAAATGTATAAAGCAG	180
Db	41	ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAla	60
QY	181	GCTTCTCTCTGACAGCAGAGAGGTGCACTTATCAAAGGATCTCCCTCAGTGGAA	240
Db	61	AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrPhe	80
QY	241	CTTAAGCAGATGAGAAGTACTCTCATCTCATCTTTCAGGCTTTTTCAGCCAGTGAT	300
Db	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSer	100
QY	301	GGAATTTGTAATGAAATTTGGTGAGGCTTTAGTCAGAGGTGCGAGTTCAGAGGCT	360
Db	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGlu	120
QY	361	CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATGTTTCACTCAGAGATGTAC	420
Db	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrPhe	140
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTATGCAATTTGA	480
Db	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIle	160
QY	481	ACCATGCCCTATGTTTAAAGAAAAGCAGATTGGGCTTCGAGTGTAGTACAGATAG	540
Db	161	ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArg	180
QY	541	TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAGAGAGGATTTTCTTCTCAGGA	600
Db	181	SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSer	200
QY	601	TCTTTTGTCTATATTTGGCTTAAAGAGAGAGTCTTATCCAGGACTCACTTTTTC	660
Db	201	SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPhe	220
QY	661	AATGAACTCATCAGCAGATGAGGACTTCACTGTGCTTTCAGTGTGCTGATGTTCCA	720
Db	221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPhe	240
QY	721	TACTTAGTAAATAAGCCTTCAGAAAGAAAGGCTCAGGAGATCATTTGTGTGTCTCA	780
Db	241	TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys	260

```
QY 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTGCATT 840
Db |||||||
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db |||||||
QY 841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGTGAACTTGGATTCTCA 900
Db |||||||
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGGCAGAAAATCCTTTTGTGATTTTATGGAACATTTCTTTAGAAGGAAA 960
Db |||||||
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGCTTATGCGAGAAACC 1020
Db |||||||
QY 321 ThrAsnPheGluLysArgValSerGluTyGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGTTCTTCCCTTGGATGCAGATTTT 1053
Db |||||||
QY 341 ThrAsnValPheThrLeuAspAlaAspPhe 351
```

Search completed: October 30, 2005, 06:50:27
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2005, 00:02:51 ; Search time 1372.5 Seconds
(without alignments)
11030.942 Million cell updates/sec

Title: US-10-698-228-2
Perfect score: 1053
Sequence: 1 atggggcaccgaaagcc.....tcacctgtgacagatttt 1053

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27723106 seqs, 7188968421 residues
Total number of hits satisfying chosen parameters: 55446212

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA, New:
1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:
2: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq2:
3: /cgn2_6/ptodata/1/pna/US05_NEW_COMB.seq:
4: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq:
5: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:
6: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:
7: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq1:
8: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq2:
9: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
10: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq1:
11: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq2:
12: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq3:
13: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq4:
14: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq5:
15: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq6:
16: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq7:
17: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq8:
18: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq9:
19: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
20: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq1:
21: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq2:
22: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq3:
23: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq4:
24: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq5:
25: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq6:
26: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq7:
27: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq8:
28: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1053	100.0	1056	1	PCT-US05-09639-16
2	1053	100.0	1056	27	US-11-088-686-16
3	1053	100.0	4929	1	PCT-US05-22501-1545
4	1053	100.0	4955	1	PCT-US05-09639-44
5	1053	100.0	4955	1	PCT-US05-21650-13

6	1053	100.0	4955	27	US-11-088-686-44	Sequence 44, Appl
7	1052.6	100.0	1849	15	US-10-990-328A-49	Sequence 49, Appl
8	1051.4	99.8	1056	1	PCT-US05-09639-18	Sequence 18, Appl
9	1051.4	99.8	1056	27	US-11-088-686-18	Sequence 18, Appl
10	1049.8	99.7	1056	1	PCT-US05-09639-20	Sequence 20, Appl
11	1049.8	99.7	1056	27	US-11-088-686-20	Sequence 20, Appl
12	1048.2	99.5	1056	1	PCT-US05-09639-22	Sequence 22, Appl
13	1048.2	99.5	1056	1	PCT-US05-09639-24	Sequence 24, Appl
14	1048.2	99.5	1056	27	US-11-088-686-22	Sequence 22, Appl
15	1048.2	99.5	1056	27	US-11-088-686-24	Sequence 24, Appl
16	1046.6	99.4	1056	1	PCT-US05-09639-26	Sequence 26, Appl
17	1046.6	99.4	1056	1	PCT-US05-09639-28	Sequence 26, Appl
18	1046.6	99.4	1056	27	US-11-088-686-26	Sequence 26, Appl
19	1046.6	99.4	1056	27	US-11-088-686-28	Sequence 28, Appl
20	624	59.3	624	1	PCT-US05-09639-42	Sequence 42, Appl
21	624	59.3	624	27	US-11-088-686-42	Sequence 42, Appl
22	578.4	54.9	1170	1	PCT-US05-09639-2	Sequence 2, Appl
23	578.4	54.9	1170	27	US-11-088-686-2	Sequence 2, Appl
24	578.4	54.9	1649	9	US-10-553-098-583	Sequence 583, App
25	578.4	54.9	1649	13	US-10-955-054A-135	Sequence 135, App
26	578.4	54.9	2470	1	PCT-US05-13260-66	Sequence 66, Appl
27	578.4	54.9	2470	1	PCT-US05-20840-66	Sequence 66, Appl
28	578.4	54.9	2470	1	PCT-US05-22501-6487	Sequence 6487, Ap
29	578.4	54.9	2479	15	US-10-940-774A-2025	Sequence 2025, Ap
30	578.4	54.9	2482	11	US-10-302-689A-107255	Sequence 107255,
31	578.4	54.9	2500	1	PCT-US05-09639-43	Sequence 43, Appl
32	578.4	54.9	2500	1	PCT-US05-20435-12	Sequence 12, Appl
33	578.4	54.9	2500	1	PCT-US05-21650-12	Sequence 12, Appl
34	578.4	54.9	2500	1	PCT-US05-27243-78	Sequence 78, Appl
35	578.4	54.9	2500	14	US-10-960-414-154	Sequence 154, App
36	578.4	54.9	2500	15	US-10-940-774A-145	Sequence 145, App
37	578.4	54.9	2500	24	US-11-000-688A-1211	Sequence 1211, Ap
38	578.4	54.9	2500	24	US-11-150-888-12	Sequence 12, Appl
39	578.4	54.9	2500	27	US-11-088-686-43	Sequence 43, Appl
40	578	54.9	2909	15	US-10-990-328A-4852	Sequence 4852, Ap
41	578	54.9	3616	15	US-10-990-328A-4851	Sequence 4851, Ap
42	576.8	54.8	1170	1	PCT-US05-09639-4	Sequence 4, Appli
43	576.8	54.8	1170	27	US-11-088-686-4	Sequence 4, Appli
44	575.2	54.6	1170	1	PCT-US05-09639-6	Sequence 6, Appli
45	575.2	54.6	1170	1	PCT-US05-09639-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
PCT-US05-09639-16
; Sequence 16, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1056)
PCT-US05-09639-16

Query Match 100.0%; Score 1053; DB 1; Length 1056;
Best Local Similarity 100.0%; Pred. No. 7,1e-271;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGCGACCCGAAAGCGGAGCGCGCGCTGATCAGGTGAGAGATCATCTTCA 60
|||||


```
Db 721 TACCTAGTAATAAGCCTTCAGAGAAAGGGTCAGGGAGATCATTTGTTGATGCTGTCAA 780
QY 781 ATTGAGCAGAGATTTTAAACAGAGCCTTGCAGATGGCTCATTTGGAATGAATTCGATT 840
Db 781 ATTGAGCAGAGATTTTAAACAGAGCCTTGCAGATGGCTCATTTGGAATGAATTCGATT 840
QY 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGTGGAACTTGGATTTCA 900
Db 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGTGGAACTTGGATTTCA 900
QY 901 AAGGTTTTTCAGGAGAGAAATCCTTTTGATTTTATGAGAAACATTTCTTTAGAGGAAA 960
Db 901 AAGGTTTTTCAGGAGAGAAATCCTTTTGATTTTATGAGAAACATTTCTTTAGAGGAAA 960
QY 961 ACAAAATTTCTTGAAGAACAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1020
Db 961 ACAAAATTTCTTGAAGAACAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1020
QY 1021 ACAGATAACGTTCTTACCTTTGATGACAGATTTT 1053
Db 1021 ACAGATAACGTTCTTACCTTTGATGACAGATTTT 1053

RESULT 3
PCT-US05-22501-1545
; Sequence 1545, Application PC/TUS0522501
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals
; TITLE OF INVENTION: Determining Cancer-Linked Gene and Therapeutic Targets Using
; FILE REFERENCE: Molecular Cytogenetic Methods
; CURRENT APPLICATION NUMBER: PCT/US05/22501
; PRIOR FILING DATE: 2005-07-07
; PRIOR APPLICATION NUMBER: 60/581,699
; PRIOR FILING DATE: 2004-06-23
; NUMBER OF SEQ ID NOS: 7840
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1545
; LENGTH: 4929
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: cDNA sequence
PCT-US05-22501-1545

Query Match 100.0%; Score 1053; DB 1; Length 4929;
Best Local Similarity 100.0%; Pred. No. 1.1e-270;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGACCCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 245 ATGGCGGACCCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 305 GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
QY 121 CGTTTGTTCATCTTTCCATCCAGTACCTGATATTTGGAAATGTATATAACAGGCACAG 180
Db 365 CGTTTGTTCATCTTTCCATCCAGTACCTGATATTTGGAAATGTATATAACAGGCACAG 424
QY 181 GCTTCCTTCTGACAGCAGAGAGGTGCACTTATCAAGAGATTCCTCCCTCACTGGAAACAG 240
Db 425 GCTTCCTTCTGACAGCAGAGAGGTGCACTTATCAAGAGATTCCTCCCTCACTGGAAACAG 484
QY 241 CTTAAGCAGATGAGAGATCTTCACTCTCTACATCTTAGCTTTTTCAGCCAGTGTAT 300
Db 485 CTTAAGCAGATGAGAGATCTTCACTCTCTACATCTTAGCTTTTTCAGCCAGTGTAT 544
QY 301 GGAATTTGTAATGAAAATTTTGGTGAGCGCTTTAGTCAGGAGGTGTCAGGTTCCAGAGCT 360
Db 545 GGAATTTGTAATGAAAATTTTGGTGAGCGCTTTAGTCAGGAGGTGTCAGGTTCCAGAGCT 604
```

```
RESULT 4
PCT-US05-09639-44
; Sequence 44, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-44

Query Match 100.0%; Score 1053; DB 1; Length 4955;
Best Local Similarity 100.0%; Pred. No. 1.1e-270;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG 664
QY 421 CTGATAGACACTTATCATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 665 CTGATAGACACTTATCATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTGGATGGATAGCAGATAGAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTGGATGGATAGCAGATAGAAA 784
QY 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGCTGTAAGAGAGTGTTCCTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGCTGTAAGAGAGTGTTCCTCTCAGGA 844
QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACATTTTCC 660
Db 845 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACATTTTCC 904
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 964
QY 721 TACTTAGTAATAAGCCTTCAGAAAGAGGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db 965 TACTTAGTAATAAGCCTTCAGAAAGAGGGTCAGGAGATCATTTGTTGATGCTGTCAA 1024
QY 781 ATTGAGCAGAGATTTTAAACAGAGCCTTGCAGATTTGGCTCATTTGSAATGAATTCATT 840
Db 1025 ATTGAGCAGAGATTTTAAACAGAGCCTTGCAGATTTGGCTCATTTGSAATGAATTCATT 1084
QY 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGTGAACTTGGATTTCA 900
Db 1085 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGTGAACTTGGATTTCA 1144
QY 901 AAGGTTTTTCAGGAGAGAAATCCTTTGATTTTATGAAAACATTTCTTTAGAGGAAA 960
Db 1145 AAGGTTTTTCAGGAGAGAAATCCTTTGATTTTATGAAAACATTTCTTTAGAGGAAA 1204
QY 961 ACAAAATTTCTTGAAGAACAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1020
Db 1205 ACAAAATTTCTTGAAGAACAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1264
QY 1021 ACAGATAACGTTCTTACCTTTGATGACAGATTTT 1053
Db 1265 ACAGATAACGTTCTTACCTTTGATGACAGATTTT 1297
```


Db 965 TACTTAGTAATAAGCTTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAAA 1024
QY 781 ATTGACGAGAGTTTTTAACAGAGAGCTTGGCAGTTGGCTCATTTGGAATGAATTCGATT 840
Db 1025 ATTGACGAGAGTTTTTAACAGAGAGCTTGGCAGTTGGCTCATTTGGAATGAATTCGATT 1084
QY 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTTGTGGAACCTTTGGATTCTCA 900
Db 1085 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTTGTGGAACCTTTGGATTCTCA 1144
QY 901 AAGGTTTTTCAGGAGAGAAATCCTTTTGAATTTATGGAAGAAACATTTCTTTAGAGAGAAA 960
Db 1145 AAGGTTTTTCAGGAGAGAAATCCTTTTGAATTTATGGAAGAAACATTTCTTTAGAGAGAAA 1204
QY 961 ACAAAATTTCTTCAGAGAGAGTTTCAGAGTATCAGCGTTTGGACGTTATGGCAGAGAAC 1020
Db 1205 ACAAAATTTCTTCAGAGAGAGTTTCAGAGTATCAGCGTTTGGACGTTATGGCAGAGAAC 1264
QY 1021 ACAGATAAAGCTTTCACCTTTGGATGAGATTTT 1053
Db 1265 ACAGATAAAGCTTTCACCTTTGGATGAGATTTT 1297

RESULT 6

US-11-088-686-44

; Sequence 44, Application US/11088686

; GENERAL INFORMATION:

; APPLICANT: Yen, Yun

; TITLE OF INVENTION: DRUG SCREENING

; FILE REFERENCE: 14037-004001

; CURRENT APPLICATION NUMBER: US/11/088,686

; CURRENT FILING DATE: 2005-03-23

; PRIOR APPLICATION NUMBER: US 60/556,836

; PRIOR FILING DATE: 2004-03-25

; NUMBER OF SEQ ID NOS: 59

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 44

; LENGTH: 4955

; TYPE: DNA

; ORGANISM: Homo sapiens

US-11-088-686-44

Query Match 100.0%; Score 1053; DB 27; Length 4955;
Best Local Similarity 100.0%; Pred. No. 1.1e-270;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGGCGGAAGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Db 245 ATGGGCGACCCGGAAGGCGGAGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 304
QY 61 GACACCAAGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 305 GACACCAAGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
QY 121 CGTTTGTTCATCTTCCATCCAGTACCTGTATTTGGAATGTATATAACAGGACAG 180
Db 365 CGTTTGTTCATCTTCCATCCAGTACCTGTATTTGGAATGTATATAACAGGACAG 424
QY 181 GCTTCTCTTGACAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAAACAG 240
Db 425 GCTTCTCTTGACAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAAACAG 484
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCTCACTCTAGCCCTTTTTCAGCCAGTGTAT 300
Db 485 CTTAAAGCAGATGAGAAGTACTTCTCTCACTCTAGCCCTTTTTCAGCCAGTGTAT 544
QY 301 GGAATGTAAATGAATAATTTGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Db 545 GGAATGTAAATGAATAATTTGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 604
QY 361 CGCTGTTCTATGGCTTTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTCTATGGCTTTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 664

QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTAATGCAATTGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTAATGCAATTGAA 724
QY 481 ACCATGCCCTATGTTTAAGAAAAAGCAGATTGGGCCCTTGGCATGATAGCAGATAGAAA 540
Db 725 ACCATGCCCTATGTTTAAGAAAAAGCAGATTGGGCCCTTGGCATGATAGCAGATAGAAA 784
QY 541 TCTACTTTGGGGAAGAGTGGTGGCCCTTGTCTGCTGTAGAGAGTTTCTTCTCAGGA 600
Db 785 TCTACTTTGGGGAAGAGTGGTGGCCCTTGTCTGCTGTAGAGAGTTTCTTCTCAGGA 844
QY 601 TCTTTTGTCTGTATATTCTGGCTAAAGAGAGGCTTATGCCAGGACTCACATTTTTC 660
Db 845 TCTTTTGTCTGTATATTCTGGCTAAAGAGAGGCTTATGCCAGGACTCACATTTTTC 904
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGTCTGCTCATGTTCCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGTCTGCTCATGTTCCAA 964
QY 721 TACTTAGTAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAAA 780
Db 965 TACTTAGTAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAAA 1024
QY 781 ATTGACGAGAGTTTTTAACAGAGAGCTTGGCAGTTGGCTCATTTGGAATGAATTCGATT 840
Db 1025 ATTGACGAGAGTTTTTAACAGAGAGCTTGGCAGTTGGCTCATTTGGAATGAATTCGATT 1084
QY 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTTGTGGAACCTTTGGATTCTCA 900
Db 1085 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTTGTGGAACCTTTGGATTCTCA 1144
QY 901 AAGGTTTTTCAGGAGAGAAATCCTTTTGAATTTATGGAAGAAACATTTCTTTAGAGAGAAA 960
Db 1145 AAGGTTTTTCAGGAGAGAAATCCTTTTGAATTTATGGAAGAAACATTTCTTTAGAGAGAAA 1204
QY 961 ACAAAATTTCTTCAGAGAAACGAGTTTCAGAGTATCAGCGTTTGGACGTTATGGCAGAGAAC 1020
Db 1205 ACAAAATTTCTTCAGAGAAACGAGTTTCAGAGTATCAGCGTTTGGACGTTATGGCAGAGAAC 1264
QY 1021 ACAGATAAAGCTTTCACCTTTGGATGAGATTTT 1053
Db 1265 ACAGATAAAGCTTTCACCTTTGGATGAGATTTT 1297

RESULT 7

US-10-990-328A-49

; Sequence 49, Application US/10990328A

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele

; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES

; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND

; TITLE OF INVENTION: USES THEREOF

; FILE REFERENCE: CL001495

; CURRENT APPLICATION NUMBER: US/10/990,328A

; CURRENT FILING DATE: 2004-11-17

; NUMBER OF SEQ ID NOS: 558824

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 49

; LENGTH: 1849

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-990-328A-49

Query Match

Best Local Similarity 100.0%; Score 1052.6; DB 15; Length 1849;

Matches 1052; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGGCGGAAGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Db 85 ATGGGCGACCCGGAAGGCGGAGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 144

QY 61 GACACAAAGTGAATAAGTCAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 120
Db |||||
145 GACACAAAGTGAATAAGTCAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 204
QY 121 CGGTTGTCTATCTTCCATCCAGTACCTCTGATATTTGGAAATGTATAAAGGACAG 180
Db |||||
205 CGGTTGTCTATCTTCCATCCAGTACCTCTGATATTTGGAAATGTATAAAGGACAG 264
QY 181 GCTTCTCTTGACAGCAGAGAGGTCGACTTATCAAGAGTCTCCTCACCTGGAACAAG 240
Db |||||
265 GCTTCTCTTGACAGCAGAGAGGTCGACTTATCAAGAGTCTCCTCACCTGGAACAAG 324
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTTAGCCTTTTTCAGCCAGTGTAT 300
Db |||||
325 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTTAGCCTTTTTCAGCCAGTGTAT 384
QY 301 GGAATTTGAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGCT 360
Db |||||
385 GGAATTTGAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGCT 444
QY 361 CGCTGTTCTATGGCTTTTCAAAATCTCATCGAATGTTCACTCAGAGATGTACAGTTTG 420
Db |||||
445 CGCTGTTCTATGGCTTTTCAAAATCTCATCGAATGTTCACTCAGAGATGTACAGTTTG 504
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db |||||
505 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 564
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTCGATGAGATGACAGATAGAAA 540
Db |||||
565 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTCGATGAGATGACAGATAGAAA 624
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAAGAGGAGTTTCTTCTCAGGA 600
Db |||||
625 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAAGAGGAGTTTCTTCTCAGGA 684
QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Db |||||
685 TCTTTTGTGCTATATTTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 744
QY 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTTGGCTTGGCTGATGTTCCAA 720
Db |||||
745 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTTGGCTTGGCTGATGTTCCAA 804
QY 721 TACTTAGTAAATAAGCCTTCAAGAAAGAGGTCAGGAGATCATTTGATGCTGTCAAA 780
Db |||||
805 TACTTAGTAAATAAGCCTTCAAGAAAGAGGTCAGGAGATCATTTGATGCTGTCAAA 864
QY 781 ATTGAGCAGAGTCTTAAAGAAAGCCTTGGCCAGTTGGCTCATTTGGAATGAATTGCATT 840
Db |||||
865 ATTGAGCAGAGTCTTAAAGAAAGCCTTGGCCAGTTGGCTCATTTGGAATGAATTGCATT 924
QY 841 TTGATGAAACAGTACATTTAGTTGTAGTGAAGATTAATCTTGGAACCTTGGAATTTCA 900
Db |||||
925 TTGATGAAACAGTACATTTAGTTGTAGTGAAGATTAATCTTGGAACCTTGGAATTTCA 984
QY 901 AAGGTTTTTCAGGCAGAAATCCTTTGATTTTATGGAATAATTTCTTTAGAGGAAAA 960
Db |||||
985 AAGGTTTTTCAGGCAGAAATCCTTTGATTTTATGGAATAATTTCTTTAGAGGAAAA 1044
QY 961 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTTATGCGAGAAACC 1020
Db |||||
1045 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTTATGCGAGAAACC 1104
QY 1021 ACAGATAACGCTTTCACCTTGGATCGAGATTTT 1053
Db |||||
1105 ACAGATAACGCTTTCACCTTGGATCGAGATTTT 1137

RESULT 8
PCT-US05-09639-18
; Sequence 18, Application PC/TUS0509639
; GENERAL INFORMATION:

; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-18

Query Match 99.8%; Score 1051.4; DB 1; Length 1056;
Best Local Similarity 99.9%; Pred. No. 1.9e-270;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGGCGGAGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db |||||
1 ATGGGCGACCCGGAAGGCGGAGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACAAAGTGAATAAGTCAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 120
Db |||||
61 GACACAAAGTGAATAAGTCAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 120
QY 121 CGGTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAATGTATAAAGGACAG 180
Db |||||
121 CGGTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAATGTATAAAGGACAG 180
QY 181 GCTTCTCTTGACAGCAGAGAGGTCGACTTATCAAGAGTCTCCTCAGGAAACAAG 240
Db |||||
181 GCTTCTCTTGACAGCAGAGAGGTCGACTTATCAAGAGTCTCCTCAGGAAACAAG 240
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTTTAGCCTTTTTCAGCCAGTGTAT 300
Db |||||
241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTTTAGCCTTTTTCAGCCAGTGTAT 300
QY 301 GGAATTTGAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGCT 360
Db |||||
301 GGAATTTGAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGCT 360
QY 361 CGCTGTTCTATGGCTTTTCAAAATCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG 420
Db |||||
361 CGCTGTTCTATGGCTTTTCAAAATCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db |||||
421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTCGATGAGATGAGAGATAGAAA 540
Db |||||
481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTCGATGAGATGAGAGATAGAAA 540
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAAGAGAGTTTCTTCTCAGGA 600
Db |||||
541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAAGAGAGTTTCTTCTCAGGA 600
QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Db |||||
601 TCTTTTGTGCTATATTTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
QY 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGAATTTGCTTGGCTGATGTTCCAA 720
Db |||||
661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGAATTTGCTTGGCTGATGTTCCAA 720
QY 721 TACTTAGTAAATAAGCCTTCAAGAAAGAGGTCAGGAGATCATTTGATGCTGTCAAA 780
Db |||||
721 TACTTAGTAAATAAGCCTTCAAGAAAGAGGTCAGGAGATCATTTGATGCTGTCAAA 780
QY 781 ATTGAGCAGAGTCTTAAAGAAAGCCTTGGCCAGTTGGCTCATTTGGAATGAATTGCATT 840
Db |||||

Db 781 ATTGAGCAGGAGTTTTTAAACAGAGCCTTCCAGTTGGCCTCAATGGAATGAATTGCATT 840
Qy 841 TTGATGAACAGTACATTTAGTTTGTAGCTGACAGATTACTGTGGAACCTTTGGAATCTCA 900
Db 841 TTGATGAACAGTACATTTAGTTTGTAGCTGACAGATTACTGTGGAACCTTTGGAATCTCA 900
Qy 901 AAGGTTTTTCAGCAGAGAAATCCTTTTGAATTTTATGGAAGAAATTTCTTTTGAAGGAAA 960
Db 901 AAGGTTTTTCAGCAGAGAAATCCTTTTGAATTTTATGGAAGAAATTTCTTTTGAAGGAAA 960
Qy 961 ACAAATTTCTTTCAGAGAACAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
Db 961 ACAAATTTCTTTCAGAGAACAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
Qy 1021 ACAGATAACGTTCTTACCTTTGGATGCAGATTTT 1053
Db 1021 ACAGATAACGTTCTTACCTTTGGATGCAGATTTT 1053

RESULT 9
US-11-088-686-18
; Sequence 18, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; PRIOR FILING DATE: 2005-03-23
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-088-686-18

Query Match 99.8%; Score 1051.4; DB 27; Length 1056;
Best Local Similarity 99.9%; Pred. No. 1.9e-270;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGGGCCCGGAAAGCGCGGAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
Db 1 ATGGGGCCCGGAAAGCGCGGAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
Qy 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Qy 121 CGGTTTGTCACTTTTCCAAATCCAGTACCTGTATTTGGAATGATTAACACAGGCAAG 180
Db 121 CGGTTTGTCACTTTTCCAAATCCAGTACCTGTATTTGGAATGATTAACACAGGCAAG 180
Qy 181 GCTTCCTCTTGACACGAGAGGTTGCACTTCAAGGNTCTCCTCACTGGACAAAG 240
Db 181 GCTTCCTCTTGACACGAGAGGTTGCACTTCAAGGNTCTCCTCACTGGACAAAG 240
Qy 241 CTTAAAGCAGATGAGAGTACTTCACTCTCACATCTTACGCTTTTTCAGCGCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAGTACTTCACTCTCACATCTTACGCTTTTTCAGCGCAGTGAT 300
Qy 301 GGAATTTGTAATGAAATTTGTGGAGCGCTTTAGTCAGAGGTGAGTTTCCAGAGGCT 360
Db 301 GGAATTTGTAATGAAATTTGTGGAGCGCTTTAGTCAGAGGTGAGTTTCCAGAGGCT 360
Qy 361 CGCTGTTTCTATGCGCTTTCAATTTCTATCGAGAAATGTTCACTCAGAGTGTACAGTTTG 420
Db 361 CGCTGTTTCTATGCGCTTTCAATTTCTATCGAGAAATGTTCACTCAGAGTGTACAGTTTG 420
Qy 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTTAAATGCAATTGAA 480
Db 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTTAAATGCAATTGAA 480

Qy 481 ACCATGCCCTATGTTTAAGAAAAAGCAGATTGGGCCCTTGGATGATAGCAGATAGAAAA 540
Db 481 ACCATGCCCTATGTTTAAGAAAAAGCAGATTGGGCCCTTGGATGATAGCAGATAGAAAA 540
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCCTTTGCTGCTGTAGAAAGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCCTTTGCTGCTGTAGAAAGAGTTTCTTCTCAGGA 600
Qy 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGGTTCTTATGCCAGGACTCCTTTTTC 660
Db 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGGTTCTTATGCCAGGACTCCTTTTTC 660
Qy 661 AATGAATCTATCAGCAGAGATGAAGGACTTCTCCTGTGCTTGTGCTGTGCTGTGCTCA 720
Db 661 AATGAATCTATCAGCAGAGATGAAGGACTTCTCCTGTGCTTGTGCTGTGCTGTGCTCA 720
Qy 721 TACTTAGTAAATAAGCCTTTCAGAGAAAGGCTCAGGAGAGATCATTTGTGTGCTGTCA 780
Db 721 TACTTAGTAAATAAGCCTTTCAGAGAAAGGCTCAGGAGAGATCATTTGTGTGCTGTCA 780
Qy 781 ATTGAGCAGGAGTTTTTAAACAGAGCCTTGGCAGTTGGCCTCATTTGGAATGAATTGCATT 840
Db 781 ATTGAGCAGGAGTTTTTAAACAGAGCCTTGGCAGTTGGCCTCATTTGGAATGAATTGCATT 840
Qy 841 TTGATGAACAGTACATTTAGTTTGTAGCTGACAGATTACTGTGGAACCTTGGATTTCTCA 900
Db 841 TTGATGAACAGTACATTTAGTTTGTAGCTGACAGATTACTGTGGAACCTTGGATTTCTCA 900
Qy 901 AAGGTTTTTTCAGCAGAGAAATCCTTTTGAATTTTATGGAAGAAATTTCTTTTGAAGGAAA 960
Db 901 AAGGTTTTTTCAGCAGAGAAATCCTTTTGAATTTTATGGAAGAAATTTCTTTTGAAGGAAA 960
Qy 961 ACAAATTTCTTTCAGAGAACAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
Db 961 ACAAATTTCTTTCAGAGAACAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
Qy 1021 ACAGATAACGTTCTTACCTTTGGATGCAGATTTT 1053
Db 1021 ACAGATAACGTTCTTACCTTTGGATGCAGATTTT 1053

RESULT 10

PCT-US05-09639-20
; Sequence 20, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-20

Query Match 99.7%; Score 1049.8; DB 1; Length 1056;
Best Local Similarity 99.8%; Pred. No. 5.1e-270;
Matches 1051; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGGGCCCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
Db 1 ATGGGGCCCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
Qy 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAAGAAAGAGTTCTCGC 120

```
121 CGTTTGTGTCATCTTTCCAAATCCAGTACCCTGATATTTGGAAATATATAACAGGCACAG 180
Db      |||
121 CGTTTGTGTCATCTTTCCAAATCCAGTACCCTGATATTTGGAAATATATAACAGGCACAG 180
Qy      |||
181 GCTTCCTTCTGCACACAGAGGTCGACTTATCAAGAGTCTCCCTCACTGGAAACAAG 240
Db      |||
181 GCTTCCTTCTGCACACAGAGAGGTCGACTTATCAAGAGTCTCCCTCACTGGAAACAAG 240
Qy      |||
241 CTTAAAGCAGATCAGAAAGTACTTCACTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db      |||
241 CTTAAAGCAGATCAGAAAGTACTTCACTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 300
Qy      |||
301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db      |||
301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Qy      |||
361 CGCTGTTTCTATGGCTTTTCAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db      |||
361 CGCTGTTTCTATGGCTTTTCAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Qy      |||
421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db      |||
421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Qy      |||
481 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGGCCCTTGGATGGATAGAGATAGAAA 540
Db      |||
481 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGGCCCTTGGATGGATAGAGATAGAAA 540
Qy      |||
541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGAGTTCCTCTCAGGA 600
Db      |||
541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGAGTTCCTCTCAGGA 600
Qy      |||
601 TCTTTTGTGCTATATCTGGCTAAAGAGAGAGGTCCTATGCGCAGGACTCACTTTTTC 660
Db      |||
601 TCTTTTGTGCTATATCTGGCTAAAGAGAGAGGTCCTATGCGCAGGACTCACTTTTTC 660
Qy      |||
661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGCCTGATGTTCCAA 720
Db      |||
661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGCCTGATGTTCCAA 720
Qy      |||
721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTGAGAGATCATTTGATGCTGTCAA 780
Db      |||
721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTGAGAGATCATTTGATGCTGTCAA 780
Qy      |||
781 ATTGACGAGGTTTTTAAACAGAGCCTTCCAGTTCGCTCATTTGGAATGAAATTGCATT 840
Db      |||
781 ATTGACGAGGTTTTTAAACAGAGCCTTCCAGTTCGCTCATTTGGAATGAAATTGCATT 840
Qy      |||
841 ATTGACGAGGTTTTTAAACAGAGCCTTCCAGTTCGCTCATTTGGAATGAAATTGCATT 900
Db      |||
841 ATTGACGAGGTTTTTAAACAGAGCCTTCCAGTTCGCTCATTTGGAATGAAATTGCATT 900
Qy      |||
901 AAGGTTTTTTCAGGAGAAAATCCTTTGATTTTATGGAAAACATTTCTTTAGAAAGAAA 960
Db      |||
901 AAGGTTTTTTCAGGAGAAAATCCTTTGATTTTATGGAAAACATTTCTTTAGAAAGAAA 960
Qy      |||
961 ACAAAATTTCTTTGAGAAAACAGTTTCAGAGTATCAGGTTTTTCAGTTATGGCAGAAAC 1020
Db      |||
961 ACAAAATTTCTTTGAGAAAACAGTTTCAGAGTATCAGGTTTTTCAGTTATGGCAGAAAC 1020
Qy      |||
1021 ACAGATAACGCTCTTCACTTGGATGCAGATTTT 1053
Db      |||
1021 ACAGATAACGCTCTTCACTTGGATGCAGATTTT 1053
```

RESULT 11
US-11-088-686-20
; Sequence 20, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088, 686

```
; CURRENT FILING DATE: 2005-03-23  
; PRIOR APPLICATION NUMBER: US 60/556,836  
; PRIOR FILING DATE: 2004-03-25  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 20  
; LENGTH: 1056  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-088-686-20  
  
Query Match 99.7%; Score 1049.8; DB 27; Length 1056;  
Best Local Similarity 99.8%; Pred. No. 5.1e-270;  
Matches 1051; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
Qy      1 ATGGCGACCCCGGAAAGCGCGGCGGCTGATCAGGATGAGATCATCTTCA 60  
Db      1 ATGGCGACCCCGGAAAGCGCGGCGGCTGATCAGGATGAGATCATCTTCA 60  
Qy      61 GACACCAACGAAAGTGAATATAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGAGTTCGCG 120  
Db      61 GACACCAACGAAAGTGAATATAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGAGTTCGCG 120  
Qy      121 CGGTTTGTCACTCTTCCAATCCAGTACCTGATATTTGGAAAAATGTATAAACAGCACAG 180  
Db      121 CGGTTTGTCACTCTTCCAATCCAGTACCTGATATTTGGAAAAATGTATAAACAGCACAG 180  
Qy      181 GCTTCCTTCTGGACAGCAGAGAGGTGCTATTAAGAGATCTCCCTCACTGGAACAG 240  
Db      181 GCTTCCTTCTGGACAGCAGAGAGGTGCTATTAAGAGATCTCCCTCACTGGAACAG 240  
Qy      241 CTTAAAGCAGATCAGAAAGTACTTCACTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 300  
Db      241 CTTAAAGCAGATCAGAAAGTACTTCACTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 300  
Qy      301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360  
Db      301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360  
Qy      361 CGCTGTTTCTATGGCTTTTCAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420  
Db      361 CGCTGTTTCTATGGCTTTTCAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420  
Qy      421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480  
Db      421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480  
Qy      481 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGGCCCTTGGATGGATAGAGATAGAAA 540  
Db      481 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGGCCCTTGGATGGATAGAGATAGAAA 540  
Qy      541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTTCCTCTCAGGA 600  
Db      541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTTCCTCTCAGGA 600  
Qy      601 TCTTTTGTGCTATATCTGGCTAAAGAGAGGTCCTATGCGCAGGACTCACTTTTTC 660  
Db      601 TCTTTTGTGCTATATCTGGCTAAAGAGAGGTCCTATGCGCAGGACTCACTTTTTC 660  
Qy      661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGCCTGATGTTCCAA 720  
Db      661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGCCTGATGTTCCAA 720  
Qy      721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTGAGAGATCATTTGATGCTGTCAA 780  
Db      721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTGAGAGATCATTTGATGCTGTCAA 780  
Qy      781 ATTGACGAGGTTTTTAAACAGAGCCTTCCAGTTCGCTCATTTGGAATGAAATTGCATT 840  
Db      781 ATTGACGAGGTTTTTAAACAGAGCCTTCCAGTTCGCTCATTTGGAATGAAATTGCATT 840  
Qy      841 TTGATGAAACAGTACATTTGATGTTGTAGCTGACAGATTTACTTGTGAACTTGGATTCTCA 900  
Db      |||
```

Db 841 TTGATGAAACAGTACATGAGTTGTAGCTGACAGATTAATCTGTGAACTTGGATCTCA 900
Qy 901 AAGGTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGGAATAATTTCTTTAGAAAGAAA 960
Db 901 AAGGTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGGAATAATTTCTTTAGAAAGAAA 960
Qy 961 ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGAGTATGCGAGAAACC 1020
Db 961 ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGAGTATGCGAGAAACC 1020
Qy 1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053

RESULT 12

PCT-US05-09639-22
; Sequence 22, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-22

Query Match 99.5%; Score 1048.2; DB 1; Length 1056;
Best Local Similarity 99.7%; Pred. No. 1.4e-269;
Matches 1050; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGGGCGACCCGGAAGCGGAGCGCGCGGCTGGATCAGATCAGATCACTTCA 60
Db 1 ATGGGCGACCCGGAAGCGGAGCGCGCGGCTGGATCAGATCAGATCACTTCA 60
Qy 61 GACACCAACGAAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAAGAGTTCTCGC 120
Qy 121 CGGTTTGTCTATCTTCCATCCAGTACCCTGATATTTGGAATGTATAACAGGCACAG 180
Db 121 CGGTTTGTCTATCTTCCATCCAGTACCCTGATATTTGGAATGTATAACAGGCACAG 180
Qy 181 GCTTCCTCTTGACAGCAGAGGTCGACTTATCAAGAGTCTCCTCACTGGAACAG 240
Db 181 GCTTCCTCTTGACAGCAGAGGTCGACTTATCAAGAGTCTCCTCACTGGAACAG 240
Qy 241 CTTTAAAGCAGATGAGAAGTACTTCTCATCTCATCTTAGCCCTTTTTCAGCCAGTGT 300
Db 241 CTTTAAAGCAGATGAGAAGTACTTCTCATCTCATCTTAGCCCTTTTTCAGCCAGTGT 300
Qy 301 GGAAATGTAAATCAAAATTTGTGGAGCGCTTTAGTCAGAGGTGCGAGGTCCAGGCT 360
Db 301 GGAAATGTAAATCAAAATTTGTGGAGCGCTTTAGTCAGAGGTGCGAGGTCCAGGCT 360
Qy 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGACAGTTTG 420
Qy 421 CTGATAGACATTTACATCAGATCCCAAGAAAGGAAATTTTATTAATCAATGAA 480
Db 421 CTGATAGACATTTACATCAGATCCCAAGAAAGGAAATTTTATTAATCAATGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTCGCGATGATAGCAGATGAAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTCGCGATGATAGCAGATGAAAA 540

Qy 541 TCTACTTTTGGGAAAAAGAGTGGTGGCTTTGCTGCTGTAGAAAGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAAAGAGTGGTGGCTTTGCTGCTGTAGAAAGAGTTTCTTCTCAGGA 600
Qy 601 TCTTTTGTCTCTATATCTTGGCTAAAGAAAGAGAGGTCTTATGCCAGGACTCACATTTTCC 660
Db 601 TCTTTTGTCTCTATATCTTGGCTAAAGAAAGAGAGGTCTTATGCCAGGACTCACATTTTCC 660
Qy 661 AATGAATCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 661 AATGAATCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Qy 721 TACTTAGTAAATAAGCCTTTCAGAAAGAAAGGTTCAGGAGATCAATTTGTATGCTGTCAAA 780
Db 721 TACTTAGTAAATAAGCCTTTCAGAAAGAAAGGTTCAGGAGATCAATTTGTATGCTGTCAAA 780
Qy 781 ATTGAGCAGGAGTTTTTTAAACAGAAAGCCTTGCAGTTGGCCTCATTTGGAATGAATTCATT 840
Db 781 ATTGAGCAGGAGTTTTTTAAACAGAAAGCCTTGCAGTTGGCCTCATTTGGAATGAATTCATT 840
Qy 841 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGTGAACTTGGATTTCTCA 900
Db 841 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGTGAACTTGGATTTCTCA 900
Qy 901 AAGGTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGGAATAATTTCTTTAGAAAGAAA 960
Db 901 AAGGTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGGAATAATTTCTTTAGAAAGAAA 960
Qy 961 ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGAGTATGCGAGAAACC 1020
Db 961 ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGAGTATGCGAGAAACC 1020
Qy 1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053

RESULT 13

PCT-US05-09639-24
; Sequence 24, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-24

Query Match 99.5%; Score 1048.2; DB 1; Length 1056;
Best Local Similarity 99.7%; Pred. No. 1.4e-269;
Matches 1050; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGGGCGACCCGGAAGCGGAGCGCGGCTGGATCAGATCAGATCACTTCA 60
Db 1 ATGGGCGACCCGGAAGCGGAGCGCGGCTGGATCAGATCAGATCACTTCA 60
Qy 61 GACACCAACGAAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAAGAGTTCTCGC 120
Qy 121 CGGTTTGTCTATCTTCCAAATCCAGTACCCTGATATTTGGAATGTATAACAGGCACAG 180
Db 121 CGGTTTGTCTATCTTCCAAATCCAGTACCCTGATATTTGGAATGTATAACAGGCACAG 180


```
Db 901 AAGGTTTTTCAGCAGAGAAATCCTTTTGATTTTATGGAAGAAACATTTCTTTAGAGAGAAA 960
Qy 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGATATCAGCGTTTTCGAGTTATGCGCAGAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGATATCAGCGTTTTCGAGTTATGCGCAGAAACC 1020
Qy 1021 ACAGATAACGTCCTTACCTTTGGATGAGAGATTTT 1053
Db 1021 ACAGATAACGTCCTTACCTTTGGATGAGAGATTTT 1053

RESULT 15
US-11-088-686-24
; Sequence 24, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-088-686-24

Query Match 99.5%; Score 1048.2; DB 27; Length 1056;
Best Local Similarity 99.7%; Pred. No. 1.4e-269;
Matches 1050; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGGGCGACCCGGAAGCGCGGCTCGATCAGATCAGAGATCATCTTCA 60
Db 1 ATGGGCGACCCGGAAGCGCGGCTCGATCAGATCAGAGATCATCTTCA 60
Qy 61 GACACCAACGAAGTGAATTAAGTCAATGAAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAGTGAATTAAGTCAATGAAGCCACTCTTAAGAAAGAGTTCTCGC 120
Qy 121 CGGTTTGTCTATCTTTCCATCCAGTACCCTGATATTTGGAATGATTAACAGGCAAG 180
Db 121 CGGTTTGTCTATCTTTCCATCCAGTACCCTGATATTTGGAATGATTAACAGGCAAG 180
Qy 181 GCTTCCTCTGGAACAGAGAGTGGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
Db 181 GCTTCCTCTGGAACAGAGAGTGGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
Qy 241 CTTAAGCAGATGAGAAAGTACTTCTCATCTCAGATCTTACCTTTTTCAGCCAGTAT 300
Db 241 CTTAAGCAGATGAGAAAGTACTTCTCATCTCAGATCTTACCTTTTTCAGCCAGTAT 300
Qy 301 GGAATTTGTAATGAAATTTTGGAGCGCTTTAGTCAGAGGTGAGTTCCAGAGGCT 360
Db 301 GGAATTTGTAATGAAATTTTGGAGCGCTTTAGTCAGAGGTGAGTTCCAGAGGCT 360
Qy 361 CGCTGTTTCTATGGCTTTCAAAATTTCTATCGAGAAATGTTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAAAATTTCTATCGAGAAATGTTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACATTTACATCAGATATCCCAAGAAAGGAAATTTTATTTAATGCAATTGAA 480
Db 421 CTGATAGACATTTACATCAGATATCCCAAGAAAGGAAATTTTATTTAATGCAATTGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAGAAAGCAGATTCGGCTTCGGATGGATAGCAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAGAAAGCAGATTCGGCTTCGGATGGATAGCAGATAGAAA 540
Qy 541 TCTACTTTTGGGGAAGAGTGGCTTTTGTCTGTGTAAGAGGAGTTTCTTCTCAGCA 600
Db 541 TCTACTTTTGGGGAAGAGTGGCTTTTGTCTGTGTAAGAGGAGTTTCTTCTCAGCA 600
```

```
Qy 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGGTCTTATGCCAGGACTCATTCTTTTCC 660
Db 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCATTCTTTTCC 660
Qy 661 AATGAACTCATTCAGCAGAGATGAAGGACTTCACTGTGACTTTGTGCTGATGTTCCAA 720
Db 661 AATGAACTCATTCAGCAGAGATGAAGGACTTCACTGTGACTTTGTGCTGATGTTCCAA 720
Qy 721 TACTTAGTAATTAAGCCCTTCAGAGAAAGGTTCAGGAGATCATTTGATGCTGTCAA 780
Db 721 TACTTAGTAATTAAGCCCTTCAGAGAAAGGTTCAGGAGATCATTTGATGCTGTCAA 780
Qy 781 ATTGAGCAGAGATTTTAAACAGAGCCCTTGGCAGTTGGCCCTCATTTGGAATGAATTCATT 840
Db 781 ATTGAGCAGAGATTTTAAACAGAGCCCTTGGCAGTTGGCCCTCATTTGGAATGAATTCATT 840
Qy 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTTACTTTGGAAACATTTCTTTAGAGGAAA 900
Db 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTTACTTTGGAAACATTTCTTTAGAGGAAA 900
Qy 901 AAGGTTTTTTCAGGCGAGAAATCCTTTTGAATTTTATGGAACATTTCTTTAGAGGAAA 960
Db 901 AAGGTTTTTTCAGGCGAGAAATCCTTTTGAATTTTATGGAACATTTCTTTAGAGGAAA 960
Qy 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGAGTTATGSCAGAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGAGTTATGSCAGAAACC 1020
Qy 1021 ACAGATAACGTCCTTTCACCTTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGTCCTTTCACCTTTGGATGAGATTTT 1053
```

Search completed: October 30, 2005, 06:01:25
Job time : 1375.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2005, 23:48:01 ; Search time 3214.5 Seconds
(without alignments)
13372.738 Million cell updates/sec

Title: US-10-698-228-2
Perfect score: 1053
Sequence: 1 atggcgacccgaaagcc.....tcacctggatgcagatttt 1053

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 45554873 seqs, 20411521753 residues
Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA Main:*

1:	/cgn2_6/ptodata/1/pna/PCTUS1 COMB.seq.*
2:	/cgn2_6/ptodata/1/pna/PCTUS2 COMB.seq.*
3:	/cgn2_6/ptodata/1/pna/PCTUS COMB.seq.*
4:	/cgn2_6/ptodata/1/pna/US06 COMB.seq.*
5:	/cgn2_6/ptodata/1/pna/US07 COMB.seq.*
6:	/cgn2_6/ptodata/1/pna/US08 COMB.seq.*
7:	/cgn2_6/ptodata/1/pna/US081 COMB.seq.*
8:	/cgn2_6/ptodata/1/pna/US082 COMB.seq.*
9:	/cgn2_6/ptodata/1/pna/US083 COMB.seq.*
10:	/cgn2_6/ptodata/1/pna/US084 COMB.seq.*
11:	/cgn2_6/ptodata/1/pna/US085 COMB.seq.*
12:	/cgn2_6/ptodata/1/pna/US086 COMB.seq.*
13:	/cgn2_6/ptodata/1/pna/US087 COMB.seq.*
14:	/cgn2_6/ptodata/1/pna/US088 COMB.seq.*
15:	/cgn2_6/ptodata/1/pna/US089 COMB.seq.*
16:	/cgn2_6/ptodata/1/pna/US090 COMB.seq.*
17:	/cgn2_6/ptodata/1/pna/US091 COMB.seq.*
18:	/cgn2_6/ptodata/1/pna/US092A COMB.seq.*
19:	/cgn2_6/ptodata/1/pna/US092B COMB.seq.*
20:	/cgn2_6/ptodata/1/pna/US093A COMB.seq.*
21:	/cgn2_6/ptodata/1/pna/US093B COMB.seq.*
22:	/cgn2_6/ptodata/1/pna/US094 COMB.seq.*
23:	/cgn2_6/ptodata/1/pna/US095A COMB.seq.*
24:	/cgn2_6/ptodata/1/pna/US095B COMB.seq.*
25:	/cgn2_6/ptodata/1/pna/US095C COMB.seq.*
26:	/cgn2_6/ptodata/1/pna/US095D COMB.seq.*
27:	/cgn2_6/ptodata/1/pna/US095E COMB.seq.*
28:	/cgn2_6/ptodata/1/pna/US096B COMB.seq.*
29:	/cgn2_6/ptodata/1/pna/US096C COMB.seq.*
30:	/cgn2_6/ptodata/1/pna/US096D COMB.seq.*
31:	/cgn2_6/ptodata/1/pna/US096E COMB.seq.*
32:	/cgn2_6/ptodata/1/pna/US097A COMB.seq.*
33:	/cgn2_6/ptodata/1/pna/US097B COMB.seq.*
34:	/cgn2_6/ptodata/1/pna/US097C COMB.seq.*
35:	/cgn2_6/ptodata/1/pna/US098A COMB.seq.*
36:	/cgn2_6/ptodata/1/pna/US098B COMB.seq.*
37:	/cgn2_6/ptodata/1/pna/US098C COMB.seq.*
38:	/cgn2_6/ptodata/1/pna/US098D COMB.seq.*
39:	/cgn2_6/ptodata/1/pna/US099A COMB.seq.*
40:	/cgn2_6/ptodata/1/pna/US099B COMB.seq.*
41:	/cgn2_6/ptodata/1/pna/US099C COMB.seq.*
42:	/cgn2_6/ptodata/1/pna/US099D COMB.seq.*
43:	/cgn2_6/ptodata/1/pna/US099E COMB.seq.*

44:	/cgn2_6/ptodata/1/pna/US099F COMB.seq.*
45:	/cgn2_6/ptodata/1/pna/US099G COMB.seq.*
46:	/cgn2_6/ptodata/1/pna/US100A COMB.seq.*
47:	/cgn2_6/ptodata/1/pna/US100B COMB.seq.*
48:	/cgn2_6/ptodata/1/pna/US101A COMB.seq.*
49:	/cgn2_6/ptodata/1/pna/US101B COMB.seq.*
50:	/cgn2_6/ptodata/1/pna/US102A COMB.seq.*
51:	/cgn2_6/ptodata/1/pna/US102B COMB.seq.*
52:	/cgn2_6/ptodata/1/pna/US103A COMB.seq.*
53:	/cgn2_6/ptodata/1/pna/US103B COMB.seq.*
54:	/cgn2_6/ptodata/1/pna/US104A COMB.seq.*
55:	/cgn2_6/ptodata/1/pna/US104B COMB.seq.*
56:	/cgn2_6/ptodata/1/pna/US105A COMB.seq.*
57:	/cgn2_6/ptodata/1/pna/US105B COMB.seq.*
58:	/cgn2_6/ptodata/1/pna/US106A COMB.seq.*
59:	/cgn2_6/ptodata/1/pna/US107A COMB.seq.*
60:	/cgn2_6/ptodata/1/pna/US107B COMB.seq.*
61:	/cgn2_6/ptodata/1/pna/US107C COMB.seq.*
62:	/cgn2_6/ptodata/1/pna/US107D COMB.seq.*
63:	/cgn2_6/ptodata/1/pna/US108A COMB.seq.*
64:	/cgn2_6/ptodata/1/pna/US108B COMB.seq.*
65:	/cgn2_6/ptodata/1/pna/US109A COMB.seq.*
66:	/cgn2_6/ptodata/1/pna/US109B COMB.seq.*
67:	/cgn2_6/ptodata/1/pna/US109C COMB.seq.*
68:	/cgn2_6/ptodata/1/pna/US110 COMB.seq.*
69:	/cgn2_6/ptodata/1/pna/US6000 COMB.seq.*
70:	/cgn2_6/ptodata/1/pna/US6001 COMB.seq.*
71:	/cgn2_6/ptodata/1/pna/US6002 COMB.seq.*
72:	/cgn2_6/ptodata/1/pna/US6003 COMB.seq.*
73:	/cgn2_6/ptodata/1/pna/US6004 COMB.seq.*
74:	/cgn2_6/ptodata/1/pna/US6005 COMB.seq.*
75:	/cgn2_6/ptodata/1/pna/US6006 COMB.seq.*
76:	/cgn2_6/ptodata/1/pna/US6007 COMB.seq.*
77:	/cgn2_6/ptodata/1/pna/US6008 COMB.seq.*
78:	/cgn2_6/ptodata/1/pna/US6009 COMB.seq.*
79:	/cgn2_6/ptodata/1/pna/US6010 COMB.seq.*
80:	/cgn2_6/ptodata/1/pna/US6011 COMB.seq.*
81:	/cgn2_6/ptodata/1/pna/US6012 COMB.seq.*
82:	/cgn2_6/ptodata/1/pna/US6013 COMB.seq.*
83:	/cgn2_6/ptodata/1/pna/US6014 COMB.seq.*
84:	/cgn2_6/ptodata/1/pna/US6015 COMB.seq.*
85:	/cgn2_6/ptodata/1/pna/US6016 COMB.seq.*
86:	/cgn2_6/ptodata/1/pna/US6017 COMB.seq.*
87:	/cgn2_6/ptodata/1/pna/US6018 COMB.seq.*
88:	/cgn2_6/ptodata/1/pna/US6019 COMB.seq.*
89:	/cgn2_6/ptodata/1/pna/US6020 COMB.seq.*
90:	/cgn2_6/ptodata/1/pna/US6021 COMB.seq.*
91:	/cgn2_6/ptodata/1/pna/US6022 COMB.seq.*
92:	/cgn2_6/ptodata/1/pna/US6023A COMB.seq.*
93:	/cgn2_6/ptodata/1/pna/US6023B COMB.seq.*
94:	/cgn2_6/ptodata/1/pna/US6024 COMB.seq.*
95:	/cgn2_6/ptodata/1/pna/US6025 COMB.seq.*
96:	/cgn2_6/ptodata/1/pna/US6026 COMB.seq.*
97:	/cgn2_6/ptodata/1/pna/US6027 COMB.seq.*
98:	/cgn2_6/ptodata/1/pna/US6028 COMB.seq.*
99:	/cgn2_6/ptodata/1/pna/US6029 COMB.seq.*
100:	/cgn2_6/ptodata/1/pna/US6030 COMB.seq.*
101:	/cgn2_6/ptodata/1/pna/US6031 COMB.seq.*
102:	/cgn2_6/ptodata/1/pna/US6032 COMB.seq.*
103:	/cgn2_6/ptodata/1/pna/US6033 COMB.seq.*
104:	/cgn2_6/ptodata/1/pna/US6034 COMB.seq.*
105:	/cgn2_6/ptodata/1/pna/US6035 COMB.seq.*
106:	/cgn2_6/ptodata/1/pna/US6036 COMB.seq.*
107:	/cgn2_6/ptodata/1/pna/US6037 COMB.seq.*
108:	/cgn2_6/ptodata/1/pna/US6038 COMB.seq.*
109:	/cgn2_6/ptodata/1/pna/US6039 COMB.seq.*
110:	/cgn2_6/ptodata/1/pna/US6040 COMB.seq.*
111:	/cgn2_6/ptodata/1/pna/US6041 COMB.seq.*
112:	/cgn2_6/ptodata/1/pna/US6042 COMB.seq.*
113:	/cgn2_6/ptodata/1/pna/US6043 COMB.seq.*
114:	/cgn2_6/ptodata/1/pna/US6044 COMB.seq.*
115:	/cgn2_6/ptodata/1/pna/US6045 COMB.seq.*
116:	/cgn2_6/ptodata/1/pna/US6046 COMB.seq.*

QY 661 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
DB 661 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
QY 721 TACTTAGTAAATAAGCTTTCAGAGAAAGGCTCAGGAGATCAATTTGATGCTGTCAA 780
DB 721 TACTTAGTAAATAAGCTTTCAGAGAAAGGCTCAGGAGATCAATTTGATGCTGTCAA 780
QY 781 ATTGAGCAGAGTTTTCAGAGAGCTTGCAGAGTTGCGCTCATTCGAAATGCAATTT 840
DB 781 ATTGAGCAGAGTTTTCAGAGAGCTTGCAGAGTTGCGCTCATTCGAAATGCAATTT 840
QY 841 TTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGTGAACTTGGATTTCA 900
DB 841 TTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGTGAACTTGGATTTCA 900
QY 901 AAGTTTTTTCAGGAGAAATCTTTTGAATTTATGAAACATTTCTTTTGAAGGAAA 960
DB 901 AAGTTTTTTCAGGAGAAATCTTTTGAATTTATGAAACATTTCTTTTGAAGGAAA 960
QY 961 ACAATTTTCTTTCAGAGAGCTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
DB 961 ACAATTTTCTTTCAGAGAGCTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
QY 1021 ACAGATAACGCTCTTCACTTGTGATGAGATTTT 1053
DB 1021 ACAGATAACGCTCTTCACTTGTGATGAGATTTT 1053

RESULT 2

US-09-787-491B-28
; Sequence 28, Application US/09787491B
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; TANG, Y. Tom;
; APPLICANT: CORLEY, Neil C.; GUEGLER, Karl J.;
; APPLICANT: GORONE, Gina A.; PATTERSON, Chandra;
; APPLICANT: HILLMAN, Jennifer L.; BAUGHN, Mariah R.;
; APPLICANT: LAL, Preeti; AZIMZAI, Yalda;
; APPLICANT: YOE, Henry; YANG, Junming
; TITLE OF INVENTION: RNA-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0600 USN
; CURRENT APPLICATION NUMBER: US/09/787,491B
; PRIOR FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: PCT/US99/21688
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US 60/128,660
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: US 60/069,391
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: US 60/183,025
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: US 60/155,246
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: US 09/158,720
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PERL Program
; SEQ ID NO 28
; LENGTH: 2596
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 2604449CB1
US-09-787-491B-28

Query Match 100.0%; Score 1053; DB 34; Length 2596;
Best Local Similarity 100.0%; Pred. No. 2.9e-276;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGACCGGAAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
|||||

DB 56 ATGGCGACCGGAAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 115
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGC 120
DB 116 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGC 175
QY 121 CGTTTGTCTCATCTTTTCCAAATCCAGTACCCCTGATATTTGGAAATGTATAACAGGCACAG 180
DB 176 CGTTTGTCTCATCTTTCCNAATCCAGTACCCCTGATATTTGGAAATGTATAACAGGCACAG 235
QY 181 GCTTCCCTTTCGAGCAGCAGAGAGGTCGATTTATCAAGAGATCTCCCTCATCTGAAACAG 240
DB 236 GCTTCCCTTTCGAGCAGCAGAGAGGTCGATTTATCAAGAGATCTCCCTCATCTGAAACAG 295
QY 241 CTTAAAGCAGATCAGAGTACTTCTCTCATCATCTTAGCCTTTTTCAGCCAGTCTAT 300
DB 296 CTTAAAGCAGATCAGAGTACTTCTCTCATCATCTTAGCCTTTTTCAGCCAGTCTAT 355
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGGTGCAGGTTCCAGAGGCT 360
DB 356 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGGTGCAGGTTCCAGAGGCT 415
QY 361 CGCTGTTTCTATGCGCTTTCAAATTTCTCATCAGAGATTTCTCATCAGAGATGACAGTTTG 420
DB 416 CGCTGTTTCTATGCGCTTTCAAATTTCTCATCAGAGATTTCTCATCAGAGATGACAGTTTG 475
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTATGCAATTCAG 480
DB 476 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTATGCAATTCAG 535
QY 481 ACCATGCCCTATGTTAAGAAAGGAGATTTGGGCTTTGCGATGATGAGAGATAGAAA 540
DB 536 ACCATGCCCTATGTTAAGAAAGGAGATTTGGGCTTTGCGATGATGAGAGATAGAAA 595
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTAGAGAGAGTTTCTTCTCAGGA 600
DB 596 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTAGAGAGAGTTTCTTCTCAGGA 655
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGCTTATGCGAGAGTCACTTTTTC 660
DB 656 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGCTTATGCGAGAGTCACTTTTTC 715
QY 661 AATGAATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
DB 716 AATGAATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 775
QY 721 TACTTAGTAAATAAGGCTTTCAGAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAA 780
DB 776 TACTTAGTAAATAAGGCTTTCAGAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAA 835
QY 781 ATTGAGCAGAGTTTTCAGAGAGCTTGCAGGTTGCGCTCATTTGAAATGAAATTCAT 840
DB 836 ATTGAGCAGAGTTTTCAGAGAGCTTGCAGGTTGCGCTCATTTGAAATGAAATTCAT 895
QY 841 TTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGGAACCTTGGATTTCA 900
DB 896 TTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGGAACCTTGGATTTCA 955
QY 901 AAGTTTTTTCAGGAGAAATCTTTTGAATTTTATGAAACATTTCTTTTGAAGGAAA 960
DB 956 AAGTTTTTTCAGGAGAAATCTTTTGAATTTTATGAAACATTTCTTTTGAAGGAAA 1015
QY 961 ACAATTTTCTTTCAGAGAGCTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
DB 1016 ACAATTTTCTTTCAGAGAGCTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1075
QY 1021 ACAGATAACGCTCTTCACTTGTGATGAGATTTT 1053
DB 1076 ACAGATAACGCTCTTCACTTGTGATGAGATTTT 1108

RESULT 3

US-60-128-660-20

; Sequence 20, Application US/60128660

; GENERAL INFORMATION:

; APPLICANT: Lal, Preeti

; APPLICANT: Tang, Y. Tom

; APPLICANT: Hillman, Jennifer L.

; APPLICANT: Yue, Henry

; APPLICANT: Yang, Junming

; APPLICANT: Baughn, Mariah R.

; TITLE OF INVENTION: RNA-ASSOCIATED PROTEINS

; FILE REFERENCE: PF-0682 P

; CURRENT APPLICATION NUMBER: US/60/128,660

; CURRENT FILING DATE: 1999-04-08

; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: PERL Program

; SEQ ID NO 20

; LENGTH: 2596

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE: -

; OTHER INFORMATION: 2604449

; US-60-128-660-20

Query Match 100.0%; Score 1053; DB 81; Length 2596;

Best Local Similarity 100.0%; Pred. No. 2.9e-276;

Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGATCATCTTCA 60
DB |||||||
QY 56 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGATCATCTTCA 115
DB |||||||
QY 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB |||||||
QY 116 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 175
DB |||||||
QY 121 CGGTTTGTCACTTTCAATCCAGTACCTGATATTTGGAAATGATTAACAGGCACAG 180
DB |||||||
QY 176 CGGTTTGTCACTTTCAATCCAGTACCTGATATTTGGAAATGATTAACAGGCACAG 235
DB |||||||
QY 181 GCTTCCTTCTGCACAGACAGAGGTCGACTTATCAAGGATCTCCTCACTGGACAG 240
DB |||||||
QY 236 GCTTCCTTCTGCACAGACAGAGGTCGACTTATCAAGGATCTCCTCACTGGACAG 295
DB |||||||
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTAGCCTTTTTCGAGCCAGTGTAT 300
DB |||||||
QY 296 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTAGCCTTTTTCGAGCCAGTGTAT 355
DB |||||||
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
DB |||||||
QY 356 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 415
DB |||||||
QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
DB |||||||
QY 416 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 475
DB |||||||
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGAA 480
DB |||||||
QY 476 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGAA 535
DB |||||||
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTTCGATGATAGCAGATAGAAAA 540
DB |||||||
QY 536 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTTCGATGATAGCAGATAGAAAA 595
DB |||||||
QY 541 TCTA CTTTGGGAAAGAGTGGTGGCTTTGCTGTGTGAAGAGGATTTTCTCTCAGGA 600
DB |||||||
QY 596 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTGAAGAGGATTTTCTCTCAGGA 655
DB |||||||
QY 601 TCTTTTGTCTATATTTCTGGCTTAAGAGAGAGGTCCTTATGCCAGCTCACTTTTCC 660
DB |||||||
QY 656 TCTTTTGTCTATATTTCTGGCTTAAGAGAGAGGTCCTTATGCCAGCTCACTTTTCC 715
DB |||||||
QY 661 AATGAACATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
DB |||||||
QY 716 AATGAACATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 775
DB |||||||
```

RESULT 4

US-10-170-235-20856

; Sequence 20856, Application US/10170235

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig

; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN

; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF

; FILE REFERENCE: CL001380

; CURRENT APPLICATION NUMBER: US/10/170,235

; CURRENT FILING DATE: 2003-03-17

; NUMBER OF SEQ ID NOS: 42514

; SEQ ID NO 20856

; LENGTH: 3397

; TYPE: DNA

; ORGANISM: HUMAN

; US-10-170-235-20856

Query Match 100.0%; Score 1053; DB 49; Length 3397;

Best Local Similarity 100.0%; Pred. No. 3.2e-276;

Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGATCATCTTCA 60
DB |||||||
QY 501 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGATCATCTTCA 560
DB |||||||
QY 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB |||||||
QY 561 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 620
DB |||||||
QY 121 CGGTTTGTCACTTTCCAAATCCAGTACCTGATATTTGGAAATGATTAACAGGCACAG 180
DB |||||||
QY 621 CGGTTTGTCACTTTCCAAATCCAGTACCTGATATTTGGAAATGATTAACAGGCACAG 680
DB |||||||
QY 181 GCTTCCTTCTGCACAGACAGAGGTCGACTTATCAAGGATCTCCTCACTGGAACAG 240
DB |||||||
QY 681 GCTTCCTTCTGCACAGACAGAGGTCGACTTATCAAGGATCTCCTCACTGGAACAG 740
DB |||||||
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTAGCCTTTTTCGAGCCAGTGTAT 300
DB |||||||
QY 741 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTAGCCTTTTTCGAGCCAGTGTAT 800
DB |||||||
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
DB |||||||
QY 801 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 860
DB |||||||
QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
DB |||||||
```


Db 861 CGCTGTTCTATGGCTTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGACAGTTTG 920
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAAATGCAATTTGAA 480
Db 921 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAAATGCAATTTGAA 980
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTCGGATGGATAGCAGATAGAAAA 540
Db 981 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTCGGATGGATAGCAGATAGAAAA 1040
QY 541 TCTACTTTTGGGAAGAGTGTGGCTTTGCTGCTGTAGAAGAGTTTCTTCTCAGGA 600
Db 1041 TCTACTTTTGGGAAGAGTGTGGCTTTGCTGCTGTAGAAGAGTTTCTTCTCAGGA 1100
QY 601 TCTTTTGTCTGTATATCTTGGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Db 1101 TCTTTTGTCTGTATATCTTGGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 1160
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGTGATGTTCCAA 720
Db 1161 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGTGATGTTCCAA 1220
QY 721 TACTTAGTAATAATACCTTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 1221 TACTTAGTAATAATACCTTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 1280
QY 781 ATTGAGCAGGAGTCTTTTAAACAGAACCTTGCAGTGGCCTCATTTGGAATGCAATT 840
Db 1281 ATTGAGCAGGAGTCTTTTAAACAGAACCTTGCAGTGGCCTCATTTGGAATGCAATT 1340
QY 841 TTGATGAACACGATCATTTGAGTTTCTAGCTGACAGATTAATTTGTGAACTTGGATTTCTCA 900
Db 1341 TTGATGAACACGATCATTTGAGTTTCTAGCTGACAGATTAATTTGTGAACTTGGATTTCTCA 1400
QY 901 AAGGTTTTTCAGGCGAGAAAATCCTTTTGATTTATGGAACAATTTCTTTAGAAGGAAAA 960
Db 1401 AAGGTTTTTCAGGCGAGAAAATCCTTTTGATTTATGGAACAATTTCTTTAGAAGGAAAA 1460
QY 961 ACAAAATTTCTTGAGAACGAGTTTCAGAGTATCAGCGTTTGGAGTTATGCGAGAAACC 1020
Db 1461 ACAAAATTTCTTGAGAACGAGTTTCAGAGTATCAGCGTTTGGAGTTATGCGAGAAACC 1520
QY 1021 ACAGATAACGCTTTCACCTTGGATGCAATTTT 1053
Db 1521 ACAGATAACGCTTTCACCTTGGATGCAATTTT 1553

RESULT 5

PCT-US02-18947-71
; Sequence 71, Application PC/TUS0218947
; GENERAL INFORMATION:
; APPLICANT: Rosetta Inpharmatics
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-228
; CURRENT APPLICATION NUMBER: PCT/US02/18947
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380, 770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AB036063
; DATABASE ENTRY DATE: 2001-06-18
PCT-US02-18947-71

Query Match 100.0%; Score 1053; DB 1; Length 4955;
Best Local Similarity 100.0%; Pred. No. 3.8e-276;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGACCCGGAAAGCCGGAGCGCGGGCTGGATCAGAGATCATCTTCA 60

RESULT 6

Db 245 ATGGCGGACCCGGAAAGCCGGAGCGCGGGCTGGATCAGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATTAAGTCAATTAAGAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 305 GACACCAACGAAAGTGAATTAAGTCAATTAAGAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
QY 121 CGGTTTGTCTATCTTTCCAAATCCAGTACCCCTGATATTTGGAAAAATGTATAAACACGACAG 180
Db 365 CGGTTTGTCTATCTTTCCAAATCCAGTACCCCTGATATTTGGAAAAATGTATAAACACGACAG 424
QY 181 GCTTCCTCTTGGACAGCAGAGAGGTCGACTTATCAAAAGGATCTCCCTCACTGCGAAACAAG 240
Db 425 GCTTCCTCTTGGACAGCAGAGAGGTCGACTTATCAAAAGGATCTCCCTCACTGCGAAACAAG 484
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTAGCCCTTTTTCAGCCAGTGTAT 300
Db 485 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTAGCCCTTTTTCAGCCAGTGTAT 544
QY 301 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTCGAGGTTCCAGAGGCT 360
Db 545 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTCGAGGTTCCAGAGGCT 604
QY 361 CGCTGTTTCTATGGCTTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGACAGTTTG 420
Db 605 CGCTGTTTCTATGGCTTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGACAGTTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAAATGCAATTTGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAAATGCAATTTGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAAA 784
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGTTTCTTCTCAGGA 844
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCACATTTTTC 660
Db 845 TCTTTTGTCTATATTTCTGGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCACATTTTTC 904
QY 661 AATGAACCTCATCAGCAGAGATGAAGACCTTCACTGTGACTTTGCTTGCCTGATGTTCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGACCTTCACTGTGACTTTGCTTGCCTGATGTTCAA 964
QY 721 TACTTAGTAATAATGAAGCTTTCAGAAAGAGGTCAGGAGATCATTTGTGATGCTGTCAAA 780
Db 965 TACTTAGTAATAATGAAGCTTTCAGAAAGAGGTCAGGAGATCATTTGTGATGCTGTCAAA 1024
QY 781 ATTGAGCAGGAGTCTTTTAAACAGAAAGCCTTTGCCAGTTGGCCTCATTTGGAATGAATTCAT 840
Db 1025 ATTGAGCAGGAGTCTTTTAAACAGAAAGCCTTTGCCAGTTGGCCTCATTTGGAATGAATTCAT 1084
QY 841 TTGATGAACACGATCATTTGAGTTTGTAGCTGACAGATTAATTTGTTGGAACCTTGAATTTCTCA 900
Db 1085 TTGATGAACACGATCATTTGAGTTTGTAGCTGACAGATTAATTTGTTGGAACCTTGAATTTCTCA 1144
QY 901 AAGGTTTTTCAGGCGAGAAAATCCTTTTGATTTTATGGAACAATTTCTTTAGAGGAAAA 960
Db 1145 AAGGTTTTTCAGGCGAGAAAATCCTTTTGATTTTATGGAACAATTTCTTTAGAGGAAAA 1204
QY 961 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
Db 1205 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1264
QY 1021 ACAGATAACGCTTTCACCTTGGATGCAATTTT 1053
Db 1265 ACAGATAACGCTTTCACCTTGGATGCAATTTT 1297

```
PCT-US04-01100-71
; Sequence 71, Application PC/TUS0401100
; GENERAL INFORMATION:
; APPLICANT: Rosetta Inpharmatics LLC
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-228
; CURRENT APPLICATION NUMBER: PCT/US04/01100
; CURRENT FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: 10/342,887
; PRIOR FILING DATE: 2003-01-15
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US04-01100-71

Query Match      100.0%; Score 1053; DB 2; Length 4955;
Best Local Similarity 100.0%; Pred. No. 3.8e-276;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGACCGGAAAGCGCGGAGCGCGGCTGGATCAGATGAGAGATCATCTTCA 60
DB 245 ATGGCGGACCGGAAAGCGCGGAGCGCGGCTGGATCAGATGAGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 305 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
QY 121 CGGTTTGTCTATCTTTCCCAATCCAGTACCTGATATTTGGAAAATGTATATAACAGGACAG 180
DB 365 CGGTTTGTCTATCTTTCCCAATCCAGTACCTGATATTTGGAAAATGTATATAACAGGACAG 424
QY 181 GCTTCTCTCTGACAGACAGAGAGTTCGACTTATCAAAGGATCTCCCTCACTGGAACAAG 240
DB 425 GCTTCTCTCTGACAGACAGAGAGTTCGACTTATCAAAGGATCTCCCTCACTGGAACAAG 484
QY 241 CTATAAGCAGATGAGAAAGTACTTCTCTCATCTCTCATCTCTTAGCCTTTTTCGAGCCAGTGTAT 300
DB 485 CTATAAGCAGATGAGAAAGTACTTCTCTCATCTCTCATCTCTTAGCCTTTTTCGAGCCAGTGTAT 544

QY 301 GGAATTTGTAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
DB 545 GGAATTTGTAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 604
QY 361 CGCTGTTCTATAGGCTTTCAAATTCATCGAGAATGTTCACTCAGAGATGACAGAGTTG 420
DB 605 CGCTGTTCTATAGGCTTTCAAATTCATCGAGAATGTTCACTCAGAGATGACAGAGTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 480
DB 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTTGCGATGGATAGCAGATAGAAAA 540
DB 725 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTTGCGATGGATAGCAGATAGAAAA 784
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAAAGGAGTTTCTTCTCAGGA 600
DB 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAAAGGAGTTTCTTCTCAGGA 844
QY 601 TCTTTTGTCTATATTTCTGGCTAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
DB 845 TCTTTTGTCTATATTTCTGGCTAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 904
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 720
DB 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 964
QY 721 TACTTAGTAAATAAGCCCTTCAGAAGAAAGGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
DB 965 TACTTAGTAAATAAGCCCTTCAGAAGAAAGGGTCAGGAGATCATTTGTTGATGCTGTCAA 1024

RESULT 7
US-10-172-118-71
; Sequence 71, Application US/10172118
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AB036063
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-71

Query Match      100.0%; Score 1053; DB 49; Length 4955;
Best Local Similarity 100.0%; Pred. No. 3.8e-276;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGACCGGAAAGCGCGGAGCGCGGCTGGATCAGATGAGAGATCATCTTCA 60
DB 245 ATGGCGGACCGGAAAGCGCGGAGCGCGGCTGGATCAGATGAGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 305 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
QY 121 CGGTTTGTCTATCTTTCCCAATCCAGTACCTGATATTTGGAAAATGTATATAACAGGACAG 180
DB 365 CGGTTTGTCTATCTTTCCCAATCCAGTACCTGATATTTGGAAAATGTATATAACAGGACAG 424
QY 181 GCTTCTCTCTGACAGACAGAGAGTTCGACTTATCAAAGGATCTCCCTCACTGGAACAAG 240
DB 425 GCTTCTCTCTGACAGACAGAGAGTTCGACTTATCAAAGGATCTCCCTCACTGGAACAAG 484
QY 241 CTATAAGCAGATGAGAAAGTACTTCTCTCATCTCTCATCTCTTAGCCTTTTTCGAGCCAGTGTAT 300
DB 485 CTATAAGCAGATGAGAAAGTACTTCTCTCATCTCTCATCTCTTAGCCTTTTTCGAGCCAGTGTAT 544
```

QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db |||||
QY 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 604
Db |||||
QY 361 CGCTGTTTCTATGGCTTTTCAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db |||||
QY 605 CGCTGTTTCTATGGCTTTTCAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 480
Db |||||
QY 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGCATGGATAGCAGATAGAAA 540
Db |||||
QY 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGCATGGATAGCAGATAGAAA 784
QY 541 TCTATCTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAAGAGAGTTTCTTCTCAGGA 600
Db |||||
QY 785 TCTATCTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAAGAGAGTTTCTTCTCAGGA 844
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAGAAAGAGAGTCTTATGCCAGACTCACTTTTCC 660
Db |||||
QY 845 TCTTTTGTGCTATATTTCTGGCTTAAGAAAGAGAGTCTTATGCCAGACTCACTTTTCC 904
QY 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db |||||
QY 905 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
QY 721 TACTTAGTAAATAGCTTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db |||||
QY 965 TACTTAGTAAATAGCTTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 1024
QY 781 ATTGAGCAGGAGTTTAAACAGAGCTTGCAGTTGGCTCATTTGGAATGAATTCATTT 840
Db |||||
QY 1025 ATTGAGCAGGAGTTTAAACAGAGCTTGCAGTTGGCTCATTTGGAATGAATTCATTT 1084
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGGAATTCATTTCA 900
Db |||||
QY 1085 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGGAATTCATTTCA 1144
QY 901 AAGGTTTTTTCAGGAGAAAAATCCTTTGATTTTATGGAAGAAATTTCTTTAGAGGAAAA 960
Db |||||
QY 1145 AAGGTTTTTTCAGGAGAAAAATCCTTTGATTTTATGGAAGAAATTTCTTTAGAGGAAAA 1204
QY 961 ACAATTTCTTTGGAAGACAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1020
Db |||||
QY 1205 ACAATTTCTTTGGAAGACAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1264
QY 1021 ACAGATAAGCTCTTCACTTGGATGCAGATTTT 1053
Db |||||
QY 1265 ACAGATAAGCTCTTCACTTGGATGCAGATTTT 1297

RESULT 8

US-10-342-887-71
; Sequence 71, Application US/10342887
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linslev, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14

; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-71

Query Match 100.0%; Score 1053; DB 52; Length 4955;
Best Local Similarity 100.0%; Pred. No. 3.8e-276;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGCACCCGGAAGCGCGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db |||||
QY 245 ATGGCGCACCCGGAAGCGCGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 304
Db |||||
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCCTAAGAAAAAGATTTCTCGC 120
Db |||||
QY 305 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCCTAAGAAAAAGATTTCTCGC 364
QY 121 CGGTTTGTCTATCTTTCCAATCCAGTACCTTGATATTTGGAAAAATGTATAAACAGGACAG 180
Db |||||
QY 365 CGGTTTGTCTATCTTTCCAATCCAGTACCTTGATATTTGGAAAAATGTATAAACAGGACAG 424
QY 181 GCTTCTCTTGGACAGCAGAGAGGTGCGATTCAAAAGGATCTCCCTCACTGGACACAG 240
Db |||||
QY 425 GCTTCTCTTGGACAGCAGAGAGGTGCGATTCAAAAGGATCTCCCTCACTGGACACAG 484
QY 241 CTTAAAGCAGATCAGAGTACTTTCATCTCACAATCTTAGCCTTTTTCAGGACAGTGTAT 300
Db |||||
QY 485 CTTAAAGCAGATCAGAGTACTTTCATCTCACAATCTTAGCCTTTTTCAGGACAGTGTAT 544
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db |||||
QY 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 604
QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db |||||
QY 605 CGCTGTTTCTATGGCTTTCAAAATTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 480
Db |||||
QY 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGCATGGATAGCAGATAGAAA 540
Db |||||
QY 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGCATGGATAGCAGATAGAAA 784
QY 541 TCTATCTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAAGAGAGTTTCTTCTCAGGA 600
Db |||||
QY 785 TCTATCTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAAGAGAGTTTCTTCTCAGGA 844
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAGAAAGAGAGTCTTATGCCAGACTCACTTTTCC 660
Db |||||
QY 845 TCTTTTGTGCTATATTTCTGGCTTAAGAAAGAGAGTCTTATGCCAGACTCACTTTTCC 904
QY 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db |||||
QY 905 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
QY 721 TACTTAGTAAATAGCTTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db |||||
QY 965 TACTTAGTAAATAGCTTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 1024
QY 781 ATTGAGCAGGAGTTTAAACAGAGCTTGCAGTTGGCTCATTTGGAATGAATTCATTT 840
Db |||||
QY 1025 ATTGAGCAGGAGTTTAAACAGAGCTTGCAGTTGGCTCATTTGGAATGAATTCATTT 1084
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGGAATTCATTTCA 900
Db |||||
QY 1085 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGGAATTCATTTCA 1144

QY 901 AAGGTTTTTCAGGCGAGAAAATCCTTTTGATTTTATGGAAGAAACATTTCTTTAGAGGAAAA 960
DB 1145 AAGGTTTTTCAGGCGAGAAAATCCTTTTGATTTTATGGAAGAAACATTTCTTTAGAGGAAAA 1204
QY 961 ACAAATTTCTTGAGAAAGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1020
DB 1205 ACAAATTTCTTGAGAAAGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1264
QY 1021 ACAGATAACGCTCTTACACCTTTGGATGCGAGATTTT 1053
DB 1265 ACAGATAACGCTCTTACACCTTTGGATGCGAGATTTT 1297

RESULT 9

US-10-698-228-3
; Sequence 3, Application US/10698228
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 3
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-3

Query Match 100.0%; Score 1053; DB 58; Length 4955;
Best Local Similarity 100.0%; Pred. No. 3.8e-276;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGGACCCGGAAGGCGGAGGCGGGCTGGATCAGGATCAGAGATCATCTTCA 60
DB 245 ATGGGGGACCCGGAAGGCGGAGGCGGGCTGGATCAGGATCAGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 305 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
QY 121 CGGTTTGTCATCTTCCATCCAGTACCTGATATTTGGAATGTATAACAGGCACAG 180
DB 365 CGGTTTGTCATCTTCCATCCAGTACCTGATATTTGGAATGTATAACAGGCACAG 424
QY 181 GCTTCCTCTTGACACGACAGAGGTCGACTTATCAAGAGATCTCCCTCACTGGAACNAG 240
DB 425 GCTTCCTCTTGACACGACAGAGGTCGACTTATCAAGAGATCTCCCTCACTGGAACNAG 484
QY 241 CTTAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 300
DB 485 CTTAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 544
QY 301 GGAATTTGTAATGAATAATTTGGTGGAGCGGTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
DB 545 GGAATTTGTAATGAATAATTTGGTGGAGCGGTTTAGTCAGAGGTGAGGTTCCAGAGGCT 604
QY 361 CGCTGTTCTATGGCTTTCAAATTTCTCATCGAATGTTTCACATCAGAGATGTACAGTTTG 420
DB 605 CGCTGTTCTATGGCTTTCAAATTTCTCATCGAATGTTTCACATCAGAGATGTACAGTTTG 664
QY 421 CTGATAGACACTTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 480

DB 665 CTGATAGACACTTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 724
QY 481 ACCATGCCCTATCTTAAGAAAAAGCAGATTGGGCCCTTGGATGATAGCAGATAGAAAA 540
DB 725 ACCATGCCCTATCTTAAGAAAAAGCAGATTGGGCCCTTGGATGATAGCAGATAGAAAA 784
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAAGAGGAGTTTCTTCTCAGGA 600
DB 785 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAAGAGGAGTTTCTTCTCAGGA 844
QY 601 TCTTTTGGCTGCTATATCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACATTTTCC 660
DB 845 TCTTTTGGCTGCTATATCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACATTTTCC 904
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGGCTGCTGATGTTCAA 720
DB 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGGCTGCTGATGTTCAA 964
QY 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
DB 965 TACTTAGTAAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 1024
QY 781 ATTGACGAGAGTTTAAACAGAGCCTTCCAGTTGGCCCTCATTTGGAATGAATGCAATT 840
DB 1025 ATTGACGAGAGTTTAAACAGAGCCTTCCAGTTGGCCCTCATTTGGAATGAATGCAATT 1084
QY 841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGGAACTTTGGATTCTCA 900
DB 1085 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGGAACTTTGGATTCTCA 1144
QY 901 AAGGTTTTTCAGGCGAGAAAATCCTTTTGAATTTTATGAAAAACATTTCTTTAAGAGGAAAA 960
DB 1145 AAGGTTTTTCAGGCGAGAAAATCCTTTTGAATTTTATGAAAAACATTTCTTTAAGAGGAAAA 1204
QY 961 ACAAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGCGCAGAAACC 1020
DB 1205 ACAAATTTCTTTCAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTATGCGCAGAAACC 1264
QY 1021 ACAGATAACGCTCTTCACTTTGGATGCGAGATTTT 1053
DB 1265 ACAGATAACGCTCTTCACTTTGGATGCGAGATTTT 1297

RESULT 10

US-10-990-328-49
; Sequence 49, Application US/10990328
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 1849
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-990-328-49

Query Match 100.0%; Score 1052.6; DB 67; Length 1849;
Best Local Similarity 99.9%; Pred. No. 3.2e-276;
Matches 1052; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGGCACCCGGAAGGCGGAGGCGGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 85 ATGGGGCACCCGGAAGGCGGAGGCGGGCTGGATCAGGATGAGAGATCATCTTCA 144
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTTCTCGC 120
DB 145 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTTCTCGC 204

```
QY 121 CGGTTTGTCATCTTTCCAAATCCAGTACCCTGATATTTGGAAATGTATAAACAGGCACAG 180
Db 205 CGGTTTGTCATCTTTCCAAATCCAGTACCCTGATATTTGGAAATGTATAAACAGGCACAG 264
QY 181 GCTTCTCTTGACAGCAGAGAGTGCAGTATTAACAAGGATCTCCCTCACTGGAAACAAG 240
Db 265 GCTTCTCTTGACAGCAGAGAGTGCAGTATTAACAAGGATCTCCCTCACTGGAAACAAG 324
QY 241 CTTAAAGCAGATGAGAAGTACCTCTCATCTCATCTCTAGCCCTTTTTCGACCCAGTAT 300
Db 325 CTTAAAGCAGATGAGAAGTACCTCTCATCTCATCTCTAGCCCTTTTTCGACCCAGTAT 384
QY 301 GGAAATTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db 385 GGAAATTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 444
QY 361 CGCTGTTCTATGCGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGACAGTTTG 420
Db 445 CGCTGTTCTATGCGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGACAGTTTG 504
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGATTTTATTAATGCAATTGAA 480
Db 505 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGATTTTATTAATGCAATTGAA 564
QY 481 ACCATGCCCTATGTTAAAGAAAAGAGATTCGGGCTTGCAGTGGATAGCAGATAGAAA 540
Db 565 ACCATGCCCTATGTTAAAGAAAAGAGATTCGGGCTTGCAGTGGATAGCAGATAGAAA 624
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCGCTTTGCTGCTGTAAGAGGAGTTTCTTCTCAGGA 600
Db 625 TCTACTTTTGGGAAAGAGTGGTGGCGCTTTGCTGCTGTAAGAGGAGTTTCTTCTCAGGA 684
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGTCTTATGCGCAGACATCACTTTTCC 660
Db 685 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGTCTTATGCGCAGACATCACTTTTCC 744
QY 661 AATGAACACTCAGCAGAGATGAAGACCTTCACTGTGACCTTTGCTTGGCTGTGATGTTCAA 720
Db 745 AATGAACACTCAGCAGAGATGAAGACCTTCACTGTGACCTTTGCTTGGCTGTGATGTTCAA 804
QY 721 TACTTAGTAAATAGCCTTTTCAAGAAAGGGTTCAGGAGATCAATTTGTGATGTTCCAA 780
Db 805 TACTTAGTAAATAGCCTTTTCAAGAAAGGGTTCAGGAGATCAATTTGTGATGTTCCAA 864
QY 781 ATTGACGAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCTATTGGAATGGAATTGCAAT 840
Db 865 ATTGACGAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCTATTGGAATGGAATTGCAAT 924
QY 841 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTAATCTGTGAACTTGGATCTCA 900
Db 925 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTAATCTGTGAACTTGGATCTCA 984
QY 901 AAGGTTTTTCAGGACAGAAATCCTTTGATTTATGGAATAATTTCTTTAGAGGAAAA 960
Db 985 AAGGTTTTTCAGGACAGAAATCCTTTGATTTATGGAATAATTTCTTTAGAGGAAAA 1044
QY 961 ACAAATTTCTTTGAAAAAGAGTTTTCAGAGTATCAGCGTTTTCAGATTATGCGAAGAAC 1020
Db 1045 ACAAATTTCTTTGAAAAAGAGTTTTCAGAGTATCAGCGTTTTCAGATTATGCGAAGAAC 1104
QY 1021 ACAGATAAGCTTTCACCTTGGATGACAGATTTT 1053
Db 1105 ACAGATAAGCTTTCACCTTGGATGACAGATTTT 1137
```

RESULT 11

US-60-505-218-20
; Sequence 20, Application US/60505218
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CANCER, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001482
; CURRENT APPLICATION NUMBER: US/60/505,218
; CURRENT FILING DATE: 2003-09-24
; NUMBER OF SEQ ID NOS: 22507
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 1849
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-505-218-20

Query Match 100.0%; Score 1052.6; DB 120; Length 1849;
Best Local Similarity 99.9%; Pred. No. 3.2e-276;
Matches 1052; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGGGCGACCCGGAAAGCGCGGAGCGGCGGCTGGATCAGATGAGATCATCTTCA 60
Db 85 ATGGGGCGACCCGGAAAGCGCGGAGCGGCGGCTGGATCAGATGAGATCATCTTCA 144
QY 61 GACACCAACGAAAGTCAAAATAAAGTCAAAATGAAGAGCCACTTCTTAAGAAAGAGTTCTCGC 120
Db 145 GACACCAACGAAAGTCAAAATAAAGTCAAAATGAAGAGCCACTTCTTAAGAAAGAGTTCTCGC 204
QY 121 CGGTTTGTCTATCTTCCAAATCCAGTACCCTGATATTTGGAAAAATGTATAAACAGGCACAG 180
Db 205 CGGTTTGTCTATCTTCCAAATCCAGTACCCTGATATTTGGAAAAATGTATAAACAGGCACAG 264
QY 181 GCTTCTCTTGACAGCAGAGAGTGCAGTATTAACAAGATCTCCCTCACTGGAACAAG 240
Db 265 GCTTCTCTTGACAGCAGAGAGTGCAGTATTAACAAGATCTCCCTCACTGGAACAAG 324
QY 241 CTTAAAGCAGATGAGAAGTACCTCTCATCTCTCATCTCTAGCCCTTTTTCGACCCAGTAT 300
Db 325 CTTAAAGCAGATGAGAAGTACCTCTCATCTCATCTCTAGCCCTTTTTCGACCCAGTAT 384
QY 301 GGAAATTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db 385 GGAAATTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 444
QY 361 CGCTGTTCTATGCGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTAAGTTG 420
Db 445 CGCTGTTCTATGCGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTAAGTTG 504
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGATTTTATTAATGCAATTGAA 480
Db 505 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGATTTTATTAATGCAATTGAA 564
QY 481 ACCATGCCCTATGTTAAAGAAAAGAGATTCGGGCTTGCAGTGGATAGCAGATAGAAA 540
Db 565 ACCATGCCCTATGTTAAAGAAAAGAGATTCGGGCTTGCAGTGGATAGCAGATAGAAA 624
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCGCTTTGCTGCTGTAAGAGGAGTTTCTTCTCAGGA 600
Db 625 TCTACTTTTGGGAAAGAGTGGTGGCGCTTTGCTGCTGTAAGAGGAGTTTCTTCTCAGGA 684
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGTCTTATGCGCAGACATCACTTTTCC 660
Db 685 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGTCTTATGCGCAGACATCACTTTTCC 744
QY 661 AATGAACACTCAGCAGAGATGAAGACCTTCACTGTGACCTTTGCTTGGCTGTGATGTTCAA 720
Db 745 AATGAACACTCAGCAGAGATGAAGACCTTCACTGTGACCTTTGCTTGGCTGTGATGTTCAA 804
QY 721 TACTTAGTAAATAGCCTTTTCAAGAAAGGGTTCAGGAGATCAATTTGTGATGTTCCAA 780
Db 805 TACTTAGTAAATAGCCTTTTCAAGAAAGGGTTCAGGAGATCAATTTGTGATGTTCCAA 864
QY 781 ATTGACGAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCTATTGGAATGGAATTGCAAT 840
Db 865 ATTGACGAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCTATTGGAATGGAATTGCAAT 924
QY 841 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTAATCTGTGAACTTGGATCTCA 900
```

Db 925 TTGATGAAACAGTACATTCGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 984
QY 901 AAGGTTTTTCAGGCGAGAAATCCTTTTGTATTTATCGAAACCAATTTCTTTAGAGGAAA 960
Db 985 AAGGTTTTTCAGGCGAGAAATCCTTTTGTATTTATGGAACCAATTTCTTTAGAGGAAA 1044
QY 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTGTGCACTTATGCGAGAAAC 1020
Db 1045 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTGTGCACTTATGCGAGAAAC 1104
QY 1021 ACAGATAACGTTCTACCTTGGATGAGATTTT 1053
Db 1105 ACAGATAACGTTCTACCTTGGATGAGATTTT 1137

RESULT 12
US-60-449-629-12
; Sequence 12, Application US/60449629
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CANCER, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001449
; CURRENT APPLICATION NUMBER: US/60/449,629
; CURRENT FILING DATE: 2003-02-04
; NUMBER OF SEQ ID NOS: 2712
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 3397
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-449-629-12

Query Match 100.0%; Score 1052.6; DB 114; Length 3397;
Best Local Similarity 99.9%; Pred. No. 4.2e-276;
Matches 1052; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGAAAGCGCGGAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
Db 501 ATGGGCGACCCGAAAGCGCGGAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 560
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Db 561 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 620
QY 121 CGGTTTGTCACTTTCCAAATCCAGTACCTGTATATTTGGAAATGTATAAACAGGCACAG 180
Db 621 CGGTTTGTCACTTTCCAAATCCAGTACCTGTATATTTGGAAATGTATAAACAGGCACAG 680
QY 181 GCTTCCTCTGGACAGCAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAACAG 240
Db 681 GCTTCCTCTGGACAGCAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAACAG 740
QY 241 CTTAAAGCAGATGAGAGTACTTCACTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db 741 CTTAAAGCAGATGAGAGTACTTCACTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 800
QY 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Db 801 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 860
QY 361 CGCTGTTCTATGGCTTTCAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db 861 CGCTGTTCTATGGCTTTCAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 920
QY 421 CTGATAGACACTTACATCAGAGATCCAAAGAAAGGAAATTTTATTTAATGCAATGAA 480
Db 921 CTGATAGACACTTACATCAGAGATCCAAAGAAAGGAAATTTTATTTAATGCAATGAA 980
QY 481 ACCATGCCCTTATGTTAAGAAAAAAGCAGATTTGGGCTTTCGATGGATAGCAGATAGAAA 540
Db 981 ACCATGCCCTTATGTTAAGAAAAAAGCAGATTTGGGCTTTCGATGGATAGCAGATAGAAA 1040

QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGATTTTCTTCTCAGGA 600
Db 1041 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGATTTTCTTCTCAGGA 1100
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAAAGAGAGGTCTTATGCGAGGACTCACATTTTCC 660
Db 1101 TCTTTTGTGCTATATTTCTGGCTTAAAGAAAGAGAGGTCTTATGCGAGGACTCACATTTTCC 1160
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTTCACGTGACATTTGCTTGCCTGATGTTCCAA 720
Db 1161 AATGAACCTCATCAGCAGAGATGAAGGACTTTCACGTGACATTTGCTTGCCTGATGTTCCAA 1220
QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 1221 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 1280
QY 781 ATTGAGCAGGAGTTTTTAAACAGAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTCAT 840
Db 1281 ATTGAGCAGGAGTTTTTAAACAGAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTCAT 1340
QY 841 TTGATGAAACAGTACATTCAGGTTTGTAGCTGACAGATTAATTTGGAACCTTGGATTTCTCA 900
Db 1341 TTGATGAAACAGTACATTCAGGTTTGTAGCTGACAGATTAATTTGGAACCTTGGATTTCTCA 1400
QY 901 AAGGTTTTTTCAGGCGAGAAATCCTTTTGAATTTTATGAAACCAATTTCTTTAGAGGAAA 960
Db 1401 AAGGTTTTTTCAGGCGAGAAATCCTTTTGAATTTTATGAAACCAATTTCTTTAGAGGAAA 1460
QY 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGTTATGCGCAGAAAC 1020
Db 1461 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGTTATGCGCAGAAAC 1520
QY 1021 ACAGATAACGTTCTTACCTTGGATGAGATTTT 1053
Db 1521 ACAGATAACGTTCTTACCTTGGATGAGATTTT 1553

RESULT 13
US-10-698-228-12
; Sequence 12, Application US/10698228
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-12

Query Match 99.8%; Score 1051.4; DB 58; Length 1053;
Best Local Similarity 99.9%; Pred. No. 5.4e-276;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGGCGACCCGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120

Db 740 TACTTAGTAAATAAGCCTT CAGAAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 799
Qy 781 ATTGACGAGGTTTTTACAGAGAGCCCTTCCAGTTGGCCTCATTTGGAATGATTCATT 840
Db 800 ATTGACGAGGTTTTTAAACAGAGCCCTTCCAGTTGGCCTCATTTGGAATGATTCATT 859
Qy 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGTGAACTTTGGATTCTCA 900
Db 860 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGTGAACTTTGGATTCTCA 919
Qy 901 AAGGTTTTTCAGGAGAAAATCCTTTTGTATTTATGGAATAAATTTCTTTAGAGGAAAA 960
Db 920 AAGGTTTTTCAGGAGAAAATCCTTTTGTATTTATGGAATAAATTTCTTTAGAGGAAAA 979
Qy 961 ACAATATTTCTTTGAGAAAACAGTTTTCAGAGTATCAGCGTTTTCAGCTTATGCGAGAAAC 1020
Db 980 ACAATATTTCTTTGAGAAAACAGTTTTCAGAGTATCAGCGTTTTCAGCTTATGCGAGAAAC 1039
Qy 1021 ACAGATAAGCTCTTCACTTGGATGCGAGATTTT 1053
Db 1040 ACAGATAAGCTCTTCACTTGGATGCGAGATTTT 1072

RESULT 15

US-09-629-469A-12810
; Sequence 12810, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/09/629,469A
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12810
; LENGTH: 1601
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (40)..(1092)
US-09-629-469A-12810

Query Match 99.7%; Score 1049.8; DB 28; Length 1601;
Best Local Similarity 99.8%; Pred. No. 1.8e-275;
Matches 1051; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGGCGACCCGGAAGGCGGAGCGGCGGCTGGATCAGGATGAGATCATCTTCA 60

Db 40 ATGGGCGACCCGGAAGGCGGAGCGGCTGGATCAGGATGAGATCATCTTCA 99
Qy 61 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCCTTAAGAAAGTTCCTCGC 120
Db 100 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCCTTAAGAAAGTTCCTCGC 159
Qy 121 CGGTTGTGTCATCTTTCCAATCCAGTACCCCTCTGATATTTGGAAAATGTATAAACAGGCACAG 180
Db 160 CGGTTGTGTCATCTTTCCAATCCAGTACCCCTCTGATATTTGGAAAATGTATAAACAGGCACAG 219
Qy 181 GCTTCCTTCTTGGACAGCAGAAAGGTCGACTTATCAAAAGGATCTCCCTCACTGGAACAAG 240
Db 220 GCTTCCTTCTTGGACAGCAGAAAGGTCGACTTATCAAAAGGATCTCCCTCACTGGAACAAG 279
Qy 241 CTTAAAGCAGATCAGAAGTACTTCATCTCTCACATCTTAGCCTTTTTTGCAGCCAGTGAT 300
Db 280 CTTAAAGCAGATCAGAAGTACTTCATCTCTCACATCTTAGCCTTTTTTGCAGCCAGTGAT 339
Qy 301 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
Db 340 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 399
Qy 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAATGTTCTCATCAGAGATGTACAGTTTG 420
Db 400 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAATGTTCTCATCAGAGATGTACAGTTTG 459
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTAATGCAATTCGA 480
Db 460 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTAATGCAATTCGA 519
Qy 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCCCTTGCATGCGATAGCAGATAGAAAA 540
Db 520 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCCCTTGCATGCGATAGCAGATAGAAAA 579
Qy 541 TCTACTTTTGGGGAAGAGTGGTGGCCCTTTGCTGCTGTAGAGAGGATTTTCTTCTCAGGA 600
Db 580 TCTACTTTTGGGGAAGAGTGGTGGCCCTTTGCTGCTGTAGAGAGGATTTTCTTCTCAGGA 639
Qy 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTATGCCAGAGACTCACATTTTTC 660
Db 640 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTATGCCAGAGACTCACATTTTTC 699
Qy 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 700 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 759
Qy 721 TACTTAGTAAATAAGCCTT CAGAAGAAAGGCTCAGGAGAGATCATTTGTTGATGCTGTCAA 780
Db 760 TACTTAGTAAATAAGCCTT CAGAAGAAAGGCTCAGGAGAGATCATTTGTTGATGCTGTCAA 819
Qy 781 ATTGAGCAGGAGTTTTTAAACAGAGCCCTTGCAGTTGGCCTCATTTGGAATGAATTCATT 840
Db 820 ATTGAGCAGGAGTTTTTAAACAGAGCCCTTGCAGTTGGCCTCATTTGGAATGAATTCATT 879
Qy 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGGAACTTTGGATTCTCA 900
Db 880 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGGAACTTTGGATTCTCA 939
Qy 901 AAGGTTTTTTCAGGAGAAAATCCTTTTGTATTTATGGAATAAATTTCTTTAGAAAGGAAAA 960
Db 940 AAGGTTTTTTCAGGAGAAAATCCTTTTGTATTTATGGAATAAATTTCTTTAGAAAGGAAAA 999
Qy 961 ACAAAATTTCTTTGAGAAAACGAGTTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAAC 1020
Db 1000 ACAAAATTTCTTTGAGAAAACGAGTTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAAC 1059
Qy 1021 ACAGATAAGCTCTTCACTTGGATGCGAGATTTT 1053
Db 1060 ACAGATAAGCTCTTCACTTGGATGCGAGATTTT 1092

Search completed: October 30, 2005, 05:15:31

Job time : 3218.5 secs

.....

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 30, 2005, 07:04:13 ; Search time 2982 Seconds
(without alignments)
4805.127 Million cell updates/sec

Title: US-10-698-228-1

Perfect score: 1821

Sequence: 1 MGDPERPEAAGLDQDERSSS.....QRFVMAETTTNNVFTLDADF 351

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 45554873 seqs, 20411521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q-/cgn2_1/USPTO.spool/US10698228/runat_26102005_100715_5163/app_query.fasta_1.519
-DB=Pending Patents NA Main -QFMT=fastap -SUFFIX=p2n.rnpsm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORW=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10698228 @cgn_1_1_3250 @runat_26102005_100715_5163
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOC=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=120 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents NA Main:*

1: /cgn2_6/ptodata/1/pna/PCTUS1_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/PCTUS2_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/PCTUS3_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/PCTUS4_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/PCTUS5_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/PCTUS6_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/PCTUS7_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/PCTUS8_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/PCTUS9_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/PCTUS10_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/PCTUS11_COMB.seq.*
12: /cgn2_6/ptodata/1/pna/PCTUS12_COMB.seq.*
13: /cgn2_6/ptodata/1/pna/PCTUS13_COMB.seq.*
14: /cgn2_6/ptodata/1/pna/PCTUS14_COMB.seq.*
15: /cgn2_6/ptodata/1/pna/PCTUS15_COMB.seq.*
16: /cgn2_6/ptodata/1/pna/PCTUS16_COMB.seq.*
17: /cgn2_6/ptodata/1/pna/PCTUS17_COMB.seq.*
18: /cgn2_6/ptodata/1/pna/PCTUS18_COMB.seq.*
19: /cgn2_6/ptodata/1/pna/PCTUS19_COMB.seq.*
20: /cgn2_6/ptodata/1/pna/PCTUS20_COMB.seq.*
21: /cgn2_6/ptodata/1/pna/PCTUS21_COMB.seq.*
22: /cgn2_6/ptodata/1/pna/PCTUS22_COMB.seq.*
23: /cgn2_6/ptodata/1/pna/PCTUS23_COMB.seq.*
24: /cgn2_6/ptodata/1/pna/PCTUS24_COMB.seq.*
25: /cgn2_6/ptodata/1/pna/PCTUS25_COMB.seq.*
26: /cgn2_6/ptodata/1/pna/PCTUS26_COMB.seq.*
27: /cgn2_6/ptodata/1/pna/PCTUS27_COMB.seq.*
28: /cgn2_6/ptodata/1/pna/PCTUS28_COMB.seq.*

29: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US097C_COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US098A_COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US098B_COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US098C_COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US098D_COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US099A_COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US099B_COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US099C_COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US099D_COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US099E_COMB.seq.*
44: /cgn2_6/ptodata/1/pna/US099F_COMB.seq.*
45: /cgn2_6/ptodata/1/pna/US099G_COMB.seq.*
46: /cgn2_6/ptodata/1/pna/US100A_COMB.seq.*
47: /cgn2_6/ptodata/1/pna/US100B_COMB.seq.*
48: /cgn2_6/ptodata/1/pna/US101A_COMB.seq.*
49: /cgn2_6/ptodata/1/pna/US101B_COMB.seq.*
50: /cgn2_6/ptodata/1/pna/US102A_COMB.seq.*
51: /cgn2_6/ptodata/1/pna/US102B_COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US103A_COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US103B_COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US104A_COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US104B_COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US105A_COMB.seq.*
57: /cgn2_6/ptodata/1/pna/US105B_COMB.seq.*
58: /cgn2_6/ptodata/1/pna/US106A_COMB.seq.*
59: /cgn2_6/ptodata/1/pna/US107A_COMB.seq.*
60: /cgn2_6/ptodata/1/pna/US107B_COMB.seq.*
61: /cgn2_6/ptodata/1/pna/US107C_COMB.seq.*
62: /cgn2_6/ptodata/1/pna/US107D_COMB.seq.*
63: /cgn2_6/ptodata/1/pna/US108A_COMB.seq.*
64: /cgn2_6/ptodata/1/pna/US108B_COMB.seq.*
65: /cgn2_6/ptodata/1/pna/US109A_COMB.seq.*
66: /cgn2_6/ptodata/1/pna/US109B_COMB.seq.*
67: /cgn2_6/ptodata/1/pna/US109C_COMB.seq.*
68: /cgn2_6/ptodata/1/pna/US110_COMB.seq.*
69: /cgn2_6/ptodata/1/pna/US6000_COMB.seq.*
70: /cgn2_6/ptodata/1/pna/US6001_COMB.seq.*
71: /cgn2_6/ptodata/1/pna/US6002_COMB.seq.*
72: /cgn2_6/ptodata/1/pna/US6003_COMB.seq.*
73: /cgn2_6/ptodata/1/pna/US6004_COMB.seq.*
74: /cgn2_6/ptodata/1/pna/US6005_COMB.seq.*
75: /cgn2_6/ptodata/1/pna/US6006_COMB.seq.*
76: /cgn2_6/ptodata/1/pna/US6007_COMB.seq.*
77: /cgn2_6/ptodata/1/pna/US6008_COMB.seq.*
78: /cgn2_6/ptodata/1/pna/US6009_COMB.seq.*
79: /cgn2_6/ptodata/1/pna/US6010_COMB.seq.*
80: /cgn2_6/ptodata/1/pna/US6011_COMB.seq.*
81: /cgn2_6/ptodata/1/pna/US6012_COMB.seq.*
82: /cgn2_6/ptodata/1/pna/US6013_COMB.seq.*
83: /cgn2_6/ptodata/1/pna/US6014_COMB.seq.*
84: /cgn2_6/ptodata/1/pna/US6015_COMB.seq.*
85: /cgn2_6/ptodata/1/pna/US6016_COMB.seq.*
86: /cgn2_6/ptodata/1/pna/US6017_COMB.seq.*
87: /cgn2_6/ptodata/1/pna/US6018_COMB.seq.*
88: /cgn2_6/ptodata/1/pna/US6019_COMB.seq.*
89: /cgn2_6/ptodata/1/pna/US6020_COMB.seq.*
90: /cgn2_6/ptodata/1/pna/US6021_COMB.seq.*
91: /cgn2_6/ptodata/1/pna/US6022_COMB.seq.*
92: /cgn2_6/ptodata/1/pna/US6023A_COMB.seq.*
93: /cgn2_6/ptodata/1/pna/US6023B_COMB.seq.*
94: /cgn2_6/ptodata/1/pna/US6024_COMB.seq.*
95: /cgn2_6/ptodata/1/pna/US6025_COMB.seq.*
96: /cgn2_6/ptodata/1/pna/US6026_COMB.seq.*
97: /cgn2_6/ptodata/1/pna/US6027_COMB.seq.*
98: /cgn2_6/ptodata/1/pna/US6028_COMB.seq.*
99: /cgn2_6/ptodata/1/pna/US6029_COMB.seq.*
100: /cgn2_6/ptodata/1/pna/US6030_COMB.seq.*
101: /cgn2_6/ptodata/1/pna/US6031_COMB.seq.*

QY 121 ArgCysPheThrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTTTCTATGGCTTTCAAAATTCATCATGAGAAATGTTCACTCAGAGATGTACAGTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 480
QY 161 ThrMetProTyrValLysLysValAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
DB 481 ACCATGCCCTATGTTAAGAAAGAGAGATGGGCTTGGATGGATAGAGATAGAAA 540
QY 181 SerThrPheGlyCysArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
DB 541 TCTACTTTTGGGGAAGAGTGGTGGCTTGTCTGTAGAGAGATTTCTTCTCAGGA 600
QY 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 601 TCTTTTGTCTGTATATCTGGCTTAAAGAGAGAGGTTCTATGCTGCTGCTGCTGCTTTC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 661 AATGAACCTATCAGCAGATGAGGACTTCACTGTGACTTTGCTTGGCTGCTGCTGCTGCT 720
QY 241 TyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAlaValLys 260
DB 721 TACTTAGTAAATAGCTTCAAGAGAGAGGTCAGGAGATCATTTGATGCTGTCAAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 781 ATTGAGCAGGAGTTTTAAACAGAGGCTTGGCCCTGCTCATTTGGAATGCAATTCAT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 841 TTGATGAAACAGTACATTCAGTTGTAGCTGACAGATTAATCTTGTGAACTTGGATCTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluLys 320
DB 901 AAGGTTTTTCAGCAGAGAAATCCTTTGATTTATGGAAGAACATTTCTTTAGAGGAAA 960
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 961 ACAATTTCTTTGAGAAAGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1021 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1053

RESULT 2

US-10-698-228-12

; Sequence 12, Application US/10698228

; GENERAL INFORMATION:

; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA

; TITLE OF INVENTION: New Protein and its DNA

; FILE REFERENCE: 2619WOOP

; CURRENT APPLICATION NUMBER: US/10/698,228

; PRIOR FILING DATE: 2003-10-30

; PRIOR APPLICATION NUMBER: US/10/019,733

; PRIOR FILING DATE: 2001-12-28

; PRIOR APPLICATION NUMBER: JP 11-181131

; PRIOR FILING DATE: 1999-06-28

; PRIOR APPLICATION NUMBER: JP 11-192391

; PRIOR FILING DATE: 1999-07-06

; PRIOR APPLICATION NUMBER: JP 2000-017770

; PRIOR FILING DATE: 2000-01-21

; NUMBER OF SEQ ID NOS: 14

; SEQ ID NO 12

; LENGTH: 1053

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION:

US-10-698-228-12

Alignment Scores:

Pred. No.: 3,78e-194 Length: 1053
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 58 Gaps: 0

US-10-698-228-1 (1-351) x US-10-698-228-12 (1-1053)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
DB 1 ATGGCGCACC CGGNAAGCGCGGAGCGCGGCTGGATCAGATGAGATGAGATCACTTCA 60
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
DB 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
DB 121 CGGTTTGTCTATCTTCCAATCCAGTACCTCTCATCTCTACATCTTAGCCTTTTTCAGCCAGTGTAT 180
QY 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
DB 181 GCTTCCTTCTGGACAGCAGAGAGGTTGACTTTATCAAGAGGATCTCCCTCCTCCTGGAAACAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
DB 241 CTTTAAAGCAGATGAGAAGTACTTCTCATCTCTACATCTTAGCCTTTTTCAGCCAGTGTAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCACAGAGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATGAGAAATGTTTCTACTCAGAGATGTACAGTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 480
QY 161 ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
DB 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCTTGCATGGATGAGATAGAAAA 540
QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
DB 541 TCTACTTTTGGGGAAGAGTGGTGGGCTTTGCTCTCTAGAAAGAGTGTCTTCTCTCAGGA 600
QY 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGCTTATGCCAGGACTCCTTTTTC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 661 AATGAACCTATCAGCAGATGAGGACTTCACTGTGACTTTGCTTGTCTGCTGCTGCTGCT 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB 721 TACTTAGTAAATAGGCTTTCAGAAAGAAAGGTCAGGAGAGATCATTTGTGTGTCTGCTCAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 781 ATTGAGCAGGAGTTTTAAACAGAGGCTTCCAGTGGCTTCCCTCATTTGGAATGAAATTCAT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 841 TTGATGAAACAGTACATTCAGTTGTAGCTGACAGATTAATCTTGTGGAACCTTGGATCTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluLys 320

Db 901 AAGGTTTTTCAGCAGAAATCCTTTGATTTTATGGAACAACATTTCTTTAGAGGAAAA 960
Qy 321 ThrAsnPhelGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 961 ACAAAATTTCTTTGAGAAACAGATTTTTCAGAGTATCAGCGTTTTGACAGTTATGCGAGAAACC 1020
Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1021 ACAGATAACGCTCTTACCTTTGGATGACAGATTTT 1053

RESULT 3

US-10-698-228-4
; Sequence 4, Application US/10698228
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-4

Alignment Scores:
Pred. No.: 3,92e-194 Length: 1081
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 58 Gaps: 0

US-10-698-228-1 (1-351) x US-10-698-228-4 (1-1081)

Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 20 ATGGGCGACCCGGAAAGCGCGNAGCGCGGCTGGATCAGATCAGATCATCTTCA 79
Qy 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
Db 80 GACACCAACGAAAGTCAATTAAGTCAATGAAGAGCCATCTCTAAGAAAGAGTTCTCGC 139
Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Db 140 CGGTTTGTCTATCTTCCATTCAGTACCTCCCTGATATTTGGAAATGTATTAACAGGCACAG 199
Qy 61 AlaSerPheTrpThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db 200 GCTTCCTCTTGACACAGAAAGGTTGACTTATCAAGATCTCCTCTACTGGAAACAAG 259
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAlaSerAsp 100
Db 260 CTTAAAGCAGATGAGAAAGTACTTCACTCTCACAATCTAGCCCTTTTTCGAGCCAGTGAT 319
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 320 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTCGAGGTTCCAGAGGCT 379
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuLysGlnValHisSerGluMetTyrSerLeu 140

Db 380 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTTACAGTTTG 439
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 440 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTATTAATGCAATTGA 499
Qy 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTTrpIleAlaAspArgLys 180
Db 500 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCCTTCGATGATAGCAGATAGAAA 559
Qy 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGluLysValPhePheSerGly 200
Db 560 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGCTAGAGAGGAGTTTCTTCTCAGGA 619
Qy 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 620 TCTTTTGGCTGCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACITTTTCC 679
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 680 AATGAACATCATCAGCAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 739
Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluLeuIleValAspAlaValLys 260
Db 740 TACTTAGTAAATTAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGCAAA 799
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 800 ATTGAGCAGGAGTTTAAACAGAGCCTTGCCACTTGGCCTCATTTGGAATGAATTCATT 859
Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 860 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATTTGTTGGAATTCGATCTCA 919
Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 920 AAGGTTTTTCAGGCGAGAAATCCTTTTGAATTTTATGAAAAACAATTTCTTTAGAGGAAAA 979
Qy 321 ThrAsnPhelGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 980 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAGAAACC 1039
Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1040 ACAGATAACGCTCTTACCTTTGGATGACAGATTTT 1072

RESULT 4

US-09-787-491B-28
; Sequence 28, Application US/09787491B
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; TANG, Y. Tom;
; APPLICANT: CORLEY, Neil C.; GUEGLER, Karl J.;
; APPLICANT: GORGONE, Gina A.; PATTERSON, Chandra;
; APPLICANT: HILLMAN, Jennifer L.; BAUGHN, Mariah R.;
; APPLICANT: LAL, Preeti; AZIMZAI, Valda;
; APPLICANT: YUE, Henry; YANG, Junming
; TITLE OF INVENTION: RNA-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0600 USN
; CURRENT APPLICATION NUMBER: US/09/787,491B
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: PCT/US99/21688
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US 60/128,660
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: US 60/069,391
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: US 60/183,025
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: US 60/155,246
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: US 09/158,720
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 38


```
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 356 CGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGCTGCCAGAGGCT 415

QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 416 CGCTGTGTTCTATGGCTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 475

QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 476 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATGAA 535

QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 536 ACCATGCCCTATGTTAAGAAAAGAGATGGGCCCTTGGCTGTAGAAGGAGTTTCTTCTCAGGA 595

QY 181 SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
Db 596 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAAGGAGTTTCTTCTCAGGA 655

QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 656 TCTTTTGTGCTATATCTGGCTAAAGAAGAGGCTTATGCCAGGACTCATTCTTTC 715

QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 716 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGGCTGTATCCAA 775

QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 776 TACTTAGTAATAAGGCTTTCAGAAAAGGCTCAGGAGAGATCATTTGATGCTGTCAAA 835

QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 836 ATTGAGCAGAGTTTTTAAACAGAGGCTTGGCCAGTTGGCTCATTTGGAATGAATTCATT 895

QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 896 TTGATGAACACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATCTCA 955

QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 956 AAGGTTTTTTCAGCAGAAAATCCTTTGATTTTATGAAAAACATTTCTTTAGAAGGAAA 1015

QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 1016 ACAAATTTCTTTGAGAAAACGAGTTTCAGAGATATCAGCGTTTTCAGATTATGGCAGAAAC 1075

QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1076 ACAGATAAGCTCTTCACTTGGNTGCAGATTTT 1108

RESULT 6
US-10-170-235-20856
; Sequence 20856, Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
; FILE REFERENCE: CL001380
; CURRENT APPLICATION NUMBER: US/10/170,235
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 20856
; LENGTH: 3397
; TYPE: DNA
; ORGANISM: HUMAN
US-10-170-235-20856

Alignment Scores:
Pred. No.: 1,87e-193 Length: 3397
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
```

```
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 49 Gaps: 0

US-10-698-228-1 (1-351) x US-10-170-235-20856 (1-3397)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
Db 501 ATGGCGCACC CGGAAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 560

QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
Db 561 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAAGAGTTCTCGC 620

QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Db 621 CGGTTTGTCTATCTTTCCATCCAGTACCTTGATATTTGGAAAATGTAATAACAGGCACAG 680

QY 61 AlaSerPheTrpThrAlaGluGluValAspLysSerLysAspLeuProHisTrpAsnLys 80
Db 681 GCCTTCTTCTGGACAGCAGAAAGAGGTCCACTTATCAAGGATCTCCCTCACTGGAACAAG 740

QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAspSer 100
Db 741 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT 800

QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 801 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGTTCCAGAGGCT 860

QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 861 CGCTGTTTCTATGCTTCCAAATTCATCAGAAATGTTCACTCAGAGATGTACAGATGTG 920

QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 921 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 980

QY 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 981 ACCATGCCCTATGTTAAGAAAAGCAGATTGGGCCCTTGGATGGATGAGATAGAGAAA 1040

QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db 1041 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTAGAGAGGATTTCTTCTCAGGA 1100

QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 1101 TCTTTTGTCTATATTTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCATTCTTTTTC 1160

QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 1161 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTTTCCAA 1220

QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 1221 TACTTAGTAATAAAGCCTTCAGAAAAGGCTCAGGAGATCATTTGATGCTGTCAAA 1280

QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 1281 ATTGAGCAGGAGTTTTTAAACAGAAAGCCTTGGCAGTTGGCCTCATTTGGAATGAATTGCATT 1340

QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 1341 TTGATGAAAACAGTACATGAGTTTGTAGCTGACAGATTACTTTGTGGAACATTTCTTCA 1400

QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 1401 AAGTTTTTTCAGGCAGAAAATCCTTTTGAATTTTATGAAAACATTTCTTTAAGAGGAAA 1460

QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 1461 ACAAATTTCTTTGAGAAAACGAGTTTCAGAGTATCAGCGTTTTCAGATTATGGCAGAAACC 1520
```

QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 |||||
 Db 1521 ACAGATAACGCTCTTACCTTGGATGCAGATTTT 1553

RESULT 7

PCT-US02-18947-71
 ; Sequence 71, Application PC/TUS0218947
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosetta Inpharmatics
 ; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
 ; FILE REFERENCE: 9301-175-228
 ; CURRENT APPLICATION NUMBER: PCT/US02/18947
 ; CURRENT FILING DATE: 2002-06-14
 ; PRIOR APPLICATION NUMBER: 60/380,770
 ; PRIOR FILING DATE: 2002-05-14
 ; NUMBER OF SEQ ID NOS: 2699
 ; SEQ ID NO 71
 ; LENGTH: 4955
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; PUBLIC INFORMATION:
 ; DATABASE ACCESSION NUMBER: AB036063
 ; DATABASE ENTRY DATE: 2001-06-18
 PCT-US02-18947-71

Alignment Scores:
 Pred. No.: 3,13e-193 Length: 4955
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Db: 1 Gaps: 0

US-10-698-228-1 (1-351) x PCT-US02-18947-71 (1-4955)

QY 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
 |||||
 Db 245 ATGGGCGACCCGGAAGCGCGGCTGGATCAGATGAGATCATCTTCA 304
 QY 21 AspThrAsnGluSerGluLeuIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
 |||||
 Db 305 GACACCAACGAAAGTGAATTAAGTCAATGAAGACCACTCTTAAGAAAGAGTTCTCGC 364
 QY 41 ArgPheValIlePheProGlnThrProAspIleTrpLysMetTrpLysGlnAlaGln 60
 |||||
 Db 365 CGGTTTGTCATCTTCCATCCAGTACCTGATATTTGGAAAATGTATAAACAGGCACAG 424
 QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 |||||
 Db 425 GCTTCCTCTTGACACGACGAGGTCGACTTATCAAGAGGATCTCCCTCACTGGAAACAAG 484
 QY 81 LeuLysAlaAspGluLysTrpPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
 |||||
 Db 485 CTTAAAGCAGATGAGAGTACTTCTCATCTTAGCTTTTTCAGCCAGTAT 544
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 |||||
 Db 545 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTAGTCAGAGGTGCAGGTTCCAGAGGCT 604
 QY 121 ArgCysPheTrpGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTrpSerLeu 140
 |||||
 Db 605 CGCTGTTCTATGGCTTTCCAAATCTCATCGAGAATGTTCTCTCAGAGATGTACAGTTTG 664
 QY 141 LeuIleAspThrTrpIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 |||||
 Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 724
 QY 161 ThrMetProTrpValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 |||||
 Db 725 ACCATGCCCTATGTTAAGAAAAAAGACAGATTGGGCTTCGATGATACAGATAGAAAA 784
 QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200

Db 785 TCTACTTTTGGGAAACAGTGGTGGCTTTCTCTCTAGAGAGTCTTCTCTCAGGA 844
 QY 201 SerPheAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 |||||
 Db 845 TCTTTTGTCTGTATATTCTGGCTAAAGAAGAGAGGTCTTATGCCAGGAGCTCACTTTTTC 904
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 |||||
 Db 905 AATGAACCTATCAGCAGAGATGAAGACTTCTCCTGTGACTTTGCTGCTGATGTTCCAA 964
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 |||||
 Db 965 TACTTAGTAATAAGCCTTCAGAAAGAGGTCTAGGAGATCATTTGTTGATGCTGTCAA 1024
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 |||||
 Db 1025 ATTGAGCAGGAGTTTAAACAGAGCCTTGCCAGTTGGCCTCATTTGGAATGAATGCATT 1084
 QY 281 LeuMetLysGlnTrpIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 |||||
 Db 1085 TTGATGAACAGTACATTTAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 1144
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 |||||
 Db 1145 AAGTTTTTCAGGCAGAAATCTTTTGAATTTATGGAACATTTCTTTAGAGGAAAA 1204
 QY 321 ThrAsnPhePheGluLysArgValSerGluTrpGlnArgPheAlaValMetAlaGluThr 340
 |||||
 Db 1205 ACAATTTCTTTCAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1264
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 |||||
 Db 1265 ACAGATAACGCTCTTCACTTGGATGCAGATTTT 1297

RESULT 8

PCT-US04-01100-71
 ; Sequence 71, Application PC/TUS0401100
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosetta Inpharmatics LLC
 ; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
 ; FILE REFERENCE: 9301-188-228
 ; CURRENT APPLICATION NUMBER: PCT/US04/01100
 ; CURRENT FILING DATE: 2004-01-15
 ; PRIOR APPLICATION NUMBER: 10/342,887
 ; PRIOR FILING DATE: 2003-01-15
 ; NUMBER OF SEQ ID NOS: 2699
 ; SEQ ID NO 71
 ; LENGTH: 4955
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 PCT-US04-01100-71

Alignment Scores:
 Pred. No.: 3,13e-193 Length: 4955
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Db: 2 Gaps: 0

US-10-698-228-1 (1-351) x PCT-US04-01100-71 (1-4955)

QY 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
 |||||
 Db 245 ATGGGCGACCCGGAAGCGCGGCTGGATCAGATGAGATCATCTTCA 304
 QY 21 AspThrAsnGluSerGluLeuIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
 |||||
 Db 305 GACACCAACGAAAGTGAATTAAGTCAATGAAGACCACTCTTCAAGAAAGAGTTCTCGC 364
 QY 41 ArgPheValIlePheProGlnThrProAspIleTrpLysMetTrpLysGlnAlaGln 60
 |||||

Db 365 CGGTTTTCATCTTTCCATCCAGTACCCTGATATTTGGAAATGTATAAACAGGCACAG 424
Qy 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db 425 GCTTCCTCTTGGACAGCAGAGAGGTCGATCTCAAGAGATCTCCCTCACTCGAACAG 484
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 485 CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCCCTTTTGGCAGCAGTAT 544
Qy 101 GlyLysValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 545 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 604
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 605 CGCTGTTTCTATGGCTTTCAAAATTCATCGAAGATGTTTCACTCAGAGATGTACAGTTG 664
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheAsnAlaIleGlu 160
Db 665 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 724
Qy 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 725 ACCATGCCCTATCTTAAGAAAAAAGCAGATTTGGCCCTTGGATGATGATAGAGAAA 784
Qy 181 SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
Db 785 TCTACTTTTGGGAAAGAGTGGGCTTTGCTGTGTAGAGAGATTTTCTTCTCAGGA 844
Qy 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 845 TCTTTTGTCTATATCTGGCTTAAGAGAGAGAGTCTTATGCGAGGACTCACATTTTCC 904
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 905 AATGAACCTCATCAGCAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 964
Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 965 TACTTAGTAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTATGCTGTCAA 1024
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 1025 ATTGACGAGAGTATTAACAGAGGCTTGCAGTTGGCTCATTTGGAATGAATTCATT 1084
Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluGlyPheSer 300
Db 1085 TTGATGAAACAGTACATTTAGTTTGTAGCTGACAGATTAATTTGTGGAACCTTGGATTCTCA 1144
Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 1145 AAGGTTTTTTCAGGACAGAAAATCCTTTTGAATTTTATGGAAACATTTCTTTAGAGGAAA 1204
Qy 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 1205 ACAAAATTTCTTGAGAAACAGATTTCAGATATCAGGTTTTCAGGTTTTCAGGATGAGAAC 1264
Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1265 ACAGATAACGCTTTCACCTTTGGATGAGATTTT 1297

RESULT 9

US-10-172-118-71
; Sequence 71, Application US/10172118
; GENERAL INFORMATION:

; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc

; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AB036063
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-71

Alignment Scores:

Pred. No.: 3.13e-193 Length: 4955
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 49 Gaps: 0

US-10-698-228-1 (1-351) x US-10-172-118-71 (1-4955)

Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
Db 245 ATGGCGCACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTTCA 304
Qy 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
Db 305 GACACCAACCAAGTGAATTAAGTCAATTAAGAGCCATCTTCAAGAGAGTCTCTGC 364
Qy 41 ArgPheValIlePhePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Db 365 CGGTTTGTCACTTCCATCCAGTACCTTGATATTTGGAAATGTATAAACAGGCACAG 424
Qy 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db 425 GCTTCCTCTTGGACAGCAGAGAGGTCGATCTTATCAAGGATCTCCCTCACTGGAACAAG 484
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 485 CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCCCTTTTTCAGCAGTAT 544
Qy 101 GlyLysValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 545 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 604
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 605 CGCTGTTTCTATGGCTTTCAAAATTCATCGAAGATGTTTCACTCAGAGATGTACAGTTG 664
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheAsnAlaIleGlu 160
Db 665 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 724
Qy 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCCCTTGGATGATGATAGAGAAA 784
Qy 181 SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
Db 785 TCTACTTTTGGGAAAGAGTGGGCTTTGCTGTGTAGAGGAGTCTTATGCGAGGACTCACATTTTCC 844
Qy 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 845 TCTTTTGTCTATATCTGGCTTAAGAGAGAGAGTCTTATGCGAGGACTCACATTTTCC 904
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240

Db 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTCGATGTTCCAA 964
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluLeuValAspAlaValLys 260
Db 965 TACTTAGTAATAAGCCTTCAGAGAAAGGTCAGGAGATCATTTGATGCTGTCAAA 1024
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 1025 ATTGAGCAGGAGTATTAACAGAGGCTTGCAGTGTGGCTCATTTGGAATGAATTCATT 1084
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 1085 TTGATGAACACGTACATGAGTTGTAGCTGCAGATTAATCTTGTGAACCTTGGATTTCTCA 1144
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnLysLeuGluGlyLys 320
Db 1145 AAGTTTTTCAGCGAGAAATCTTTGATTTTATGGAAGAAACATTTCTTTAGAGGAAAA 1204
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 1205 ACAATTTCTTTGAGAAACAGGTTTCAGAGTATCAGCGTTTGTGCACTTATGGCAGAAACC 1264
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1265 ACAGATAACGCTCTTCACCTTGGATGCGATTTT 1297

RESULT 10

US-10-342-887-71
; Sequence 71, Application US/10342887
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-71

Alignment Scores:

Pred. No.: 3,136-193 Length: 4955
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 52 Gaps: 0

US-10-698-228-1 (1-351) x US-10-342-887-71 (1-4955)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 245 ATGGCGGACCCCGAAGGCGCGCGCTGGATCAGGATGAGAGATCATCTTCA 304
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
Db 305 GACACCAAGAAAGTAATGAATGAAGGCCACTCTTAAGAAAGAGTTCTCGC 364

RESULT 11

US-10-698-228-3
; Sequence 3, Application US/10698228
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30

QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Db 365 CGGTTTGTTCATCTTCCATCCAGTACCTGATATTGGAAAAATGTATAAACAGGCACAG 424
QY 61 AlaSerPheTrpThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db 425 GCTTCTCTCTGGACAGCAGAAAGGTCGACTTATCAAGAGGATCTCCCTCACTGGAAACAG 484
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAspSer 100
Db 485 CTTAAAGCAGATGAGAGTACTTCTCATCTCAGATCTTAGCCTTTTTCAGGCCAGTGAT 544
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 545 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTCGAGGTTCCAGAGGCT 604
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 605 CGCTGTTTCTATGGCTTCAAAATCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 664
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTATATGCAATTGAA 724
QY 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 725 ACCATGCCCTATGTTAAGAAAAAGAGATGGGCTTGGATGATGATGAGATAGAAAA 784
QY 181 SerThrPheGlyGluArgValAlaPheAlaValGluGluValPhePheSerGly 200
Db 785 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 844
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 845 TCTTTTGTGCTATATTCTGGCTAAAGAGAGAGGTCCTATGCGCAGGACTCACATTTTCC 904
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCA 964
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 965 TACTTAGTAATAAGCCTTCAGAAAGGTCAGGAGATCATTTGATGCTGTCAAA 1024
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 1025 ATTGAGCAGGAGTATTAACAGAGGCTTGCAGTGTGGCTCATTTGGAATGAATTCATT 1084
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 1085 TTGATGAACACGTACATGAGTTGTAGCTGCAGATTAATCTTGTGGAACCTTGGATTTCTCA 1144
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnLysSerLeuGluGlyLys 320
Db 1145 AAGTTTTTCAGCGAGAAATCTTTGATTTTATGGAAGAAACATTTCTTTAGAGGAAAA 1204
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 1205 ACAATTTCTTTGAGAAACAGGTTTCAGAGTATCAGCGTTTGTGCACTTATGGCAGAAACC 1264
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1265 ACAGATAACGCTCTTCACCTTGGATGCGATTTT 1297


```
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 3
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-3

Alignment Scores:
Pred. No.: 3,13e-193 Length: 4955
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 58 Gaps: 0

US-10-698-228-1 (1-351) x US-10-698-228-3 (1-4955)

QY 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGluAspGluArgSerSerSer 20
DB 245 ATGGGCGACCCGGAAGCGCGAAGCGCGCGGCTGGATCAGATGAGATCATCTTCA 304
QY 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 305 GACACCAACGAAGTGAATAAGTCAATGAAGAGCCACTCCTAAGAAAGAGTTCTCCG 364
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
DB 365 CGTTTGTGTCATCTTCCATCCAGTACCTGATATTTGGAAATGTATATAACAGGCACAG 424
QY 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
DB 425 GCTTCCTCTCGACAGCAGAGAGGTGCACTTATCAAGGATCTCCTCACTGGAAACAAG 484
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
DB 485 CTAAAGCAGATGAGAAGTACTTCTCTCATCTTACATCTTACCTTTTTCGAGCCAGTGAT 544
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 545 GGAATGTAAATGAAAATTTGGTGAGCGCTTGTAGTCAGAGGTGCAGGTTCCAGAGGCT 604
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 605 CGCTGTTTCTATGGCTTCAATTTCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 664
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheIleuPheAsnAlaIleGlu 160
DB 665 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAAATTTTATTTAATGCAATGTAA 724
QY 161 ThrMetProTyrValLysLysValAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
DB 725 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTCGATGGATGATCAGATAGAAA 784
QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
DB 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGGAGTTTCTTCTCAGGA 844
QY 201 SerPheAlaAlaIlePheTyrLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 845 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGAGGTCTTATGCGAGGACTCACTTTTCC 904
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 904
```

```
905 AATGAACATCATCAGCAGATGAAGGACTTCACTGTGACTTTTGCTTGCCTGATGTTCCAA 964
241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260
965 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAA 1024
261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
1025 ATTGACGAGAGTGTAAACAGAAAGCCTTGCAGATTTGGCCTCATTTGGAATGAATGCATT 1084
281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGlyPheSer 300
1085 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 1144
301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
1145 AAGGTTTTTCAGGCAGAAAATCCTTTTGATTTTATGGAACATTTCTTTAGAAAGGAAA 1204
321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
1205 ACAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGCAGAAACC 1264
341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
1265 ACAGATAAGCTCTTCACTTGGATGAGATTTT 1297

RESULT 12
US-09-629-469A-12810
; Sequence 12810, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/09/629,469A
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12810
; LENGTH: 1601
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (40)..(1092)
US-09-629-469A-12810

Alignment Scores:
Pred. No.: 1.89e-193 Length: 1601
Score: 1817.00 Matches: 350
```

Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.72%	Mismatches:	0
Query Match:	99.78%	Indels:	0
DB:	28	Gaps:	0
US-10-698-228-1 (1-351) x US-09-629-469A-12810 (1-1601)			
QY	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer	20
DB	40	ATGGGGACCCCGAAGGCGGAGCGGCGGGCTGGATCAGATCAGATCATCTTCA	99
QY	21	AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerSerArg	40
DB	100	GACACCAACGAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC	159
QY	41	ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln	60
DB	160	CGGTTTGTTCATCTTTCCATCCAGTACCTCTGATATTGGAAAATGTATAACAGGACACAG	219
QY	61	AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys	80
DB	220	GGTTCTTCTTGACAGCAGAAGGTGCTGACTTATCAAGGATCTCCCTCAGTGGAAACAG	279
QY	81	LeuLysAlaAspGluLysTyrPheLysSerHisIleLeuAlaPhePheAlaLaserAsp	100
DB	280	CTTAAAGCAGATGAGAAGTACTTCTCTCATCTTAGCCCTTTTTCGACCCAGTGTAT	339
QY	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla	120
DB	340	CGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT	399
QY	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140
DB	400	CGCTGCTTCTAGGCTTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGACAGTTTG	459
QY	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160
DB	460	CTGATAGACACTTATCATCAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA	519
QY	161	ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys	180
DB	520	ACCATGCCCTATGTTAAGAAAAGCAGATTGGGCTTCGGATGGATAGACAGATAAAAA	579
QY	181	SerThrPheGlyClnArgValAlaPheAlaAlaValGluGlyValPhePheSerGly	200
DB	580	TCTACTTTTGGCGAAGAGTGGTGGCTTGTCTGTAGAGAGATTTTCTCTCAGGA	639
QY	201	SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer	220
DB	640	TCTTTTGTCTATATTTCTGGCTAAGAAAGAGAGGTCTTATGCCAGGACTCACTTTTCC	699
QY	221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln	240
DB	700	AATGAACCTCATCAGCAGAGATCAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA	759
QY	241	TyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAlaValLys	260
DB	760	TACTTAGTAAATAAGCCCTTCGAAGAAAGGGTTCAGGGAGATCATTTGTCGTCTGCAA	819
QY	261	IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle	280
DB	820	ATTGACGAGGAGTTTAAACAGAAGCCCTTGCAGTTGGCCTCATTTGGAGTAATGCAAT	879
QY	281	LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer	300
DB	880	TTGATGAACAGTACATTTAGTTTGTAGCTGACAGATTACTTTGTGAACCTTGATTTCA	939
QY	301	LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys	320
DB	940	AAGGTTTTTTCAGGCAAGAAATCTTTTGAATTTATGGAAAAACATTTCTTTAGAGGAAA	999
QY	321	ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr	340

DB	1000	ACAAATTTCTTTTGAGAAACGAGTTTTCAGAGTATCAGCGCTTTTTCAGTTATGCGAGAAACC	1059
QY	341	ThrAspAsnValPheThrLeuAspAlaAspPhe	351
DB	1060	ACAGATAACGCTCTTCACCTTGGATGCGAGATTTT	1092
RESULT 13			
US-10-917-503-12810			
; Sequence 12810 Application US/10917503			
; GENERAL INFORMATION:			
; APPLICANT: OTA, TOSHIO			
; APPLICANT: ISOGAI, TAKAO			
; APPLICANT: NISHIKAWA, TETSUO			
; APPLICANT: HAYASHI, KOJI			
; APPLICANT: SAITO, KAORU			
; APPLICANT: YAMAMOTO, JUNICHI			
; APPLICANT: ISHII, SHIZUKO			
; APPLICANT: SUGIYAMA, TOMOYASU			
; APPLICANT: WAKAMATSU, AI			
; APPLICANT: NAGAI, KEIICHI			
; APPLICANT: OTSUKI, TETSUJI			
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE			
; FILE REFERENCE: 084335/0123			
; CURRENT APPLICATION NUMBER: US/10/917,503			
; CURRENT FILING DATE: 2004-08-13			
; PRIOR APPLICATION NUMBER: US/09/629,469			
; PRIOR FILING DATE: 2000-07-28			
; PRIOR APPLICATION NUMBER: JP 1999-248036			
; PRIOR FILING DATE: 1999-07-29			
; PRIOR APPLICATION NUMBER: JP 1999-300253			
; PRIOR FILING DATE: 1999-08-27			
; PRIOR APPLICATION NUMBER: JP 2000-118776			
; PRIOR FILING DATE: 2000-01-11			
; PRIOR APPLICATION NUMBER: JP 2000-183767			
; PRIOR FILING DATE: 2000-05-02			
; PRIOR APPLICATION NUMBER: JP 2000-241899			
; PRIOR FILING DATE: 2000-06-09			
; PRIOR APPLICATION NUMBER: 60/159,590			
; PRIOR FILING DATE: 1999-10-18			
; PRIOR APPLICATION NUMBER: 60/183,322			
; PRIOR FILING DATE: 2000-02-17			
; NUMBER OF SEQ ID NOS: 19025			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 12810			
; LENGTH: 1601			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (40)..(1092)			
US-10-917-503-12810			
Alignment Scores:			
Pred. No.:	1,89e-193	Length:	1601
Score:	1817.00	Matches:	350
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.72%	Mismatches:	0
Query Match:	99.78%	Indels:	0
DB:	65	Gaps:	0
US-10-698-228-1 (1-351) x US-10-917-503-12810 (1-1601)			
QY	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer	20
DB	40	ATGGCGACCCCGAAGGCGGAGCGGCTGGATCAGATCAGATCATCTTCA	99
QY	21	AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg	40
DB	100	GACACCAACGAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC	159
QY	41	ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln	60
DB	160	CGGTTTGTTCATCTTTCCATCCAGTACCTCTGATATTGGAAAATGTATAACAGGACACAG	219

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 1849
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-990-328-49

Alignment Scores:
Pred. No.: 3 87e-193      Length: 1849
Score: 1815.00      Matches: 350
Percent Similarity: 99.72%      Conservative: 0
Best Local Similarity: 99.72%      Mismatches: 1
Query Match: 99.67%      Indels: 0
DB: 67      Gaps: 0

US-10-698-228-1 (1-351) x US-10-990-328-49 (1-1849)

Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
Db 85 ATGGCGCACC CGCAAGGCGCGGCTGGATCAGGATCAGGATCATCTTCA 144
Qy 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
Db 145 GACACCAACCGAAAGTGAATCAAAATGAAGAGCCACTCTTAAGAAAGAGTCTCGC 204
Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
Db 205 CGGTGTGTCTATCTTCCATCCAGTACCTGATATTTGGAANAATGATAAACAGGCACAG 264
Qy 61 AlaSerPheThrPheAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
Db 265 GCITTCCTTCTGGACAGCAGAGAGGTCGACTTATCAAGAGATCTCCCTCACTGGAACAAG 324
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
Db 325 CTTAAAGCAGATGAGAAGTACTTCTCATCTCAGATCTTAGCCCTTTTTCAGCCAGTGAT 384
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 385 GGAATGTAAATGAANAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGCT 444
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 445 CGCTGTTTCTATGCTTCAAAATTCATCAGAGAAATGTTCACTCAGAGATGTCAGATTTG 504
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 505 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 564
Qy 161 ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
Db 565 ACCATGCCCTTATGTTAAGAAAAAAGCAGATTGGGCTTGGATGGATAGAGATAGAAAA 624
Qy 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db 625 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTAGAGAGAGTTCCTTCTCAGA 684
Qy 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 685 TCTTTTGTCTATATCTGGCTAAAGAGAGAGGCTCTTATGCCAGGACTCACATTTTTC 744
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 745 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGTCTGCTGATGTTCAA 804
Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 805 TACTTAGTAATTAAGCTTCAGAAAGAAAGGTCAGGAGATCATTTGTTGCTGCTCAA 864
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 865 ATTGAGCAGGAGTTTTTAAACAGAAAGCCTTGCCAGTTGGCCCTCATTTGGAATGATTCATT 924

RESULT 14
US-10-990-328-49
; Sequence 49, Application US/10990328
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
```

QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 925 TTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 984
QY 301 LysValPheGluAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 985 AAGGTTTTTCAGCAGACAAATCCTTTTGATTTTATGGAACAACATTTCTTTAGAAAGGAAA 1044
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 1045 ACAATTTCTTTCAGAAACAGATTTCAGAGTATCAGCGTTTTCAGATTATGGCAGAAACC 1104
QY 341 ThrAsnValPheThrLeuAspAlaAspPhe 351
DB 1105 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1137

RESULT 15

US-60-505-218-20
; Sequence 20, Application US/60505218
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CANCER, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001482
; CURRENT APPLICATION NUMBER: US/60/505,218
; NUMBER OF SEQ ID NOS: 22507
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 1849
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-505-218-20

Alignment Scores:

Pred. No.: 3,87e-193 Length: 1849
Score: 1815.00 Matches: 350
Percent Similarity: 99.72% Conservative: 0
Best Local Similarity: 99.72% Mismatches: 1
Query Match: 99.67% Indels: 0
DB: 120 Gaps: 0

US-10-698-228-1 (1-351) x US-60-505-218-20 (1-1849)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 85 ATGGCGACCCGGAAGGCGGAGCGCGCTGGATCAGGATGAGAGATCATCTTCA 144
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 145 GACACCAAGAAAGTGAATTAAGTCAATGAAGAGCCACTCTCAAGAAAGAGTTCTCGC 204
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
DB 205 CGSTTTGTCTATCTTCCATCCAGTACCCTGATATTGGAAATGATATAACAGGCACAG 264
QY 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
DB 265 GCTTCCTCTGGACAGCAGAGAGGTGCGACTTATCAAGAGATCTCCCTCACTGGAACAAG 324
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
DB 325 CTTAAAGCAGATGAGAAGTACTTCTCATCTCATCTTACCTTTTTCAGCCAGTGTAT 384
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 385 GGAATTGTAAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGGTTCCAGAGGCT 444
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 445 CGCTGTTTCTATGGCTTTCAATTTCTCATCGAGATGTTCACTCAGAGATGTACAGTTG 504
QY 141 LeuIleAspThrTyrIleArgaspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160

Search completed: October 30, 2005, 09:45:43

Job time : 2999 secs

DB 505 CTGATAGACACTTACATCAGGATCCCAAGAAAAGGGAATTTTATTATTAATGCAATTGAA 564
QY 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTTrpIleAlaAspArgLys 180
DB 565 ACCATGCCCTATGTATAGAAAAGAGAGATTGGGCTTGGCATGGATAGCAGATAGAAA 624
QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
DB 625 TCTACTTTTGGGGAAGAGTGGTGGCTTGTCTGTAGAAAGGAGTTTCTTCTCAGGA 684
QY 201 SerPheAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 685 TCTTTTCTGCTATATTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 744
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 745 AATGAACCTCATCAGCAGAGATGAGGACTTTCACCTGTGACTTGTGCTGTGCTGATGTTCCAA 804
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAlaAlaValLys 260
DB 805 TACTTAGTAATAAGCTTTCAGAAAGAGGTTCAGGAGATCATTTGTGATGCTGTCAAA 864
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 865 ATTGAGCAGGAGTTTAAACAGAGAGCTTGGCAGTTGGCCTCATTTGGAATGAATTGCATT 924
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 925 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTTGTGAAACTTGGATTCTCA 984
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 985 AAGTTTTTTCAGGAGAAAATCTTTTGTATTTTATGGAACAACATTTCTTTAGAAAGGAAA 1044
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 1045 ACAATTTCTTTCAGAAACAGATTTCAGAGTATCAGCGTTTTCAGATTATGGCAGAAACC 1104
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1105 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1137

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 30, 2005, 07:12:02 ; Search time 1266 Seconds
(without alignments)
3986.300 Million cell updates/sec

Title: US-10-698-228-1
Perfect score: 1821
Sequence: 1 MGDPRPEAAGLDQDERSSS.....QRFAVMAETTDNVFTLDADF 351

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27723106 seqs, 7188968421 residues

Total number of hits satisfying chosen parameters: 55446212

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10698228/runat_26102005_100716_5193/app_query.fasta_1.519
-DB=Pending_Patents_NA_New -OPMT=fastap -SUFFIX=p2n.rnpn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=0 -USER=US10698228 @CIGN 1.1 141 @runat_26102005_100716_5193
-MAXLEN=200000000 -MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCk=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending_Patents_NA_New:
1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:
2: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:
3: /cgn2_6/ptodata/1/pna/US05_NEW_COMB.seq:
4: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq:
5: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:
6: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:
7: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:
8: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:
9: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
10: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
11: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
12: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
13: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
14: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
15: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
16: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
17: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
18: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
19: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
20: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
21: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
22: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
23: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
24: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
25: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
26: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
27: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
28: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1821	100.0	1056	1	PCT-US05-09639-16	Sequence 16, Appl
2	1821	100.0	1056	27	US-11-088-686-16	Sequence 16, Appl
3	1821	100.0	4929	1	PCT-US05-22501-1545	Sequence 1545, Ap
4	1821	100.0	4955	1	PCT-US05-09639-44	Sequence 44, Appl
5	1821	100.0	4955	1	PCT-US05-21650-13	Sequence 13, Appl
6	1821	100.0	4955	27	US-11-088-686-44	Sequence 44, Appl
7	1817	99.8	1056	1	PCT-US05-09639-18	Sequence 18, Appl
8	1817	99.8	1056	27	US-11-088-686-18	Sequence 18, Appl
9	1816	99.7	1056	1	PCT-US05-09639-20	Sequence 20, Appl
10	1816	99.7	1056	27	US-11-088-686-20	Sequence 20, Appl
11	1815	99.7	1849	15	US-10-990-328A-49	Sequence 49, Appl
12	1813	99.6	1056	1	PCT-US05-09639-22	Sequence 22, Appl
13	1813	99.6	1056	27	US-11-088-686-22	Sequence 22, Appl
14	1812	99.5	1056	1	PCT-US05-09639-24	Sequence 24, Appl
15	1812	99.5	1056	1	PCT-US05-09639-26	Sequence 26, Appl
16	1812	99.5	1056	27	US-11-088-686-24	Sequence 24, Appl
17	1812	99.5	1056	27	US-11-088-686-26	Sequence 26, Appl
18	1811	99.5	1056	1	PCT-US05-09639-28	Sequence 28, Appl
19	1811	99.5	1056	27	US-11-088-686-28	Sequence 28, Appl
20	1460.5	80.2	1170	1	PCT-US05-09639-2	Sequence 2, Appli
21	1460.5	80.2	1170	27	US-11-088-686-2	Sequence 2, Appli
22	1460.5	80.2	1649	9	US-10-553-098-583	Sequence 583, App
23	1460.5	80.2	1649	13	US-10-955-054A-135	Sequence 135, App
24	1460.5	80.2	2470	1	PCT-US05-13260-66	Sequence 66, Appl
25	1460.5	80.2	2470	1	PCT-US05-20840-66	Sequence 66, Appl
26	1460.5	80.2	2470	1	PCT-US05-22501-6487	Sequence 6487, Ap
27	1460.5	80.2	2479	15	US-10-940-774A-2025	Sequence 2025, Ap
28	1460.5	80.2	2482	11	US-10-302-689A-107255	Sequence 107255,
29	1460.5	80.2	2500	1	PCT-US05-09639-43	Sequence 43, Appl
30	1460.5	80.2	2500	1	PCT-US05-20435-12	Sequence 12, Appl
31	1460.5	80.2	2500	1	PCT-US05-21650-12	Sequence 12, Appl
32	1460.5	80.2	2500	1	PCT-US05-27243-78	Sequence 78, Appl
33	1460.5	80.2	2500	14	US-10-960-414-154	Sequence 154, App
34	1460.5	80.2	2500	15	US-10-940-774A-145	Sequence 145, App
35	1460.5	80.2	2500	24	US-11-000-688A-1211	Sequence 1211, Ap
36	1460.5	80.2	2500	24	US-11-150-888-12	Sequence 12, Appl
37	1460.5	80.2	2500	27	US-11-088-686-13	Sequence 43, Appl
38	1457.5	80.0	2909	15	US-10-990-328A-4852	Sequence 4852, Ap
39	1457.5	80.0	3616	15	US-10-990-328A-4851	Sequence 4851, Ap
40	1456.5	80.0	1170	1	PCT-US05-09639-4	Sequence 4, Appli
41	1456.5	80.0	1170	27	US-11-088-686-4	Sequence 4, Appli
42	1455.5	79.9	1170	1	PCT-US05-09639-6	Sequence 6, Appli
43	1455.5	79.9	1170	27	US-11-088-686-6	Sequence 6, Appli
44	1452.5	79.8	1170	1	PCT-US05-09639-8	Sequence 8, Appli
45	1452.5	79.8	1170	27	US-11-088-686-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
PCT-US05-09639-16
; Sequence 16, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16

```
/ LENGTH: 1056
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)...(1056)
PCT-US05-09639-16

Alignment Scores:
Pred. No.: 3e-182 Length: 1056
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-10-698-228-1 (1-351) x PCT-US05-09639-16 (1-1056)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 1 ATGGCGAGCCCGAAGAGCGGAGCGCGCGGCTGGATCAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 61 GACACCAACGAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
DB 121 CGGTTTGTCTATCTTCCATCCAGTACCTCGATTTGGAAATGTATTAACAGGCACAG 180
QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
DB 181 GCTTCTCTTGGACAGCAGAGAGGTGCGATTCATCAAGAGATCTCCTCACTGGAACAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAlaSerAsp 100
DB 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCTCATCTTACATCTTGGCCTTTTTCGAGCAGTGAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuLysGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGACAGTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACACTTACATCAGATCCCAAGAAAGGAAATTTTATTTAATGCAATGAA 480
QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
DB 481 ACCATCCCTATGTTAAGAAAGAGTGGGCTTGGCGCTTGCATGGATGAGATAGAAAA 540
QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGluValPhePheSerGly 200
DB 541 TCTACTTTTGGGAAAGAGTGGTGGCTTGTGCTGTAGAGAGTGTTCCTTCTCAGGA 600
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 601 TCTTTTGTCTATATCTTGGCTAAAGAGAGAGGTCTTATGCGAGGACTCACTTTTTC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 661 AATGAACATCATCAGCAGATGAAGACITTCATGTGACTTTGCTGCTGCTGCTGCTGCT 720
QY 241 TyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAlaValLys 260
DB 721 TACTTAGTAAATGAAGCTTCAGAAAGAGAGGTGAGGAGATCATTTGTGATGCTGCAAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 781 ATTGACGAGGAGTTTAAACAGAGGCTTGGCAGTTGGCCTCATTTGAATGAATTCATT 840
```

```
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 901 AAGGTTTTTCAGGCAGAAAAATCCTTTTGAATTTATGGAACAATTTCTTTAGAGGAAAA 960
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAGAAACC 1020
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1021 ACAGATAACGTCCTTCACCTTGGATGACAGATTTT 1053
```

RESULT 2

```
US-11-088-686-16
; Sequence 16, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1056)
US-11-088-686-16
```

```
Alignment Scores:
Pred. No.: 3e-182 Length: 1056
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 27 Gaps: 0
```

US-10-698-228-1 (1-351) x US-11-088-686-16 (1-1056)

```
QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 1 ATGGCGAGCCCGAAGAGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 61 GACACCAACGAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
DB 121 CGGTTTGTCTATCTTCCATCCAGTACCTCGATTTGGAAATGTATTAACAGGCACAG 180
QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
DB 181 GCTTCTCTTGGACAGCAGAGAGGTGCGATTCATCAAGAGATCTCCTCACTGGAACAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAlaSerAsp 100
DB 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCTCATCTTACATCTTTTTCGAGCAGTGAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
```


Qy	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAenValHisSerGluMetTyrSerIleu	140
Db	361	CGCTGTTTCTATGGCTTTCAAATTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG	420
Qy	141	LeuIleAspThrTyrIleArgAspProIysIysArgGluPheLeuPheAenAlaIleGlu	160
Db	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTATTAATGCAATTCAA	480
Qy	161	ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTTrpIleAlaAspArgLys	180
Db	481	ACCATGCCCTATGTTTAGAAAAAAGCAGATGGGGCTTGCATGGATAGCAGATAGAAAA	540
Qy	181	SerThrPheGlyGluArgValAlaPheAlaIaValGluGlyValPhePheSerGly	200
Db	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGCAGATTTCTTCTCAGGA	600
Qy	201	SerPheAlaIaIlePheTrpLeuLysIysArgGlyLeuMetProGlyLeuThrPheSer	220
Db	601	TCFTTTCGTGCTATATTTCGCTAAAGAAGAGAGGTCTTATGCCCAGGACTCACATTTTTC	660
Qy	221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln	240
Db	661	AATGAACCTCATCAGCAGAGATGAAGGACTTCACATGTGACTTTGCTTGCTGATGTTCCAA	720
Qy	241	TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys	260
Db	721	TACTTAGTAATTAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCA	780
Qy	261	IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle	280
Db	781	ATTGAGCAGCAGATTTTAAACAGAAAGCCTTGCCAGTTGGCCCTCATTTGGAATGAAATGCATT	840
Qy	281	LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer	300
Db	841	TTGATGAACAGTACATTTGATTTGTAGCTGACAGATTTACTTGTGGAACTTGGATTTCTCA	900
Qy	301	LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys	320
Db	901	AAGTTTTTTCAGGCAGAAAAATCCCTTTTGATTTTATGGAACATCTTTCTTAGAAGAGAAA	960
Qy	321	ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr	340
Db	961	ACAAATTTCTTCAGAAACCGATTTACAGATATCAGCGTTTTCGCGTTATGGCAGAAACC	1020
Qy	341	ThrAspAenValPheThrLeuAspAlaAspPhe	351
Db	1021	ACAGATAACCTCTTCACCTTGGATGCAGATTTT	1053

```

RESULT 3
PCT-US05-22501-1545
; Sequence 1545, Application PC/TUS0522501
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals
; TITLE OF INVENTION: Determining Cancer-Linked Gene and Therapeutic Targets Using
; TITLE OF INVENTION: Molecular Cytogenetic Methods
; FILE REFERENCE: 689290-249
; CURRENT APPLICATION NUMBER: PCT/US05/22501
; CURRENT FILING DATE: 2005-07-07
; PRIOR APPLICATION NUMBER: 60/581,699
; PRIOR FILING DATE: 2004-06-23
; NUMBER OF SEQ ID NOS: 7840
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1545
; LENGTH: 4929
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: cdna sequence
PCT-US05-22501-1545

Alignment Scores:
Pred. No.: 2.13e-181 Length: 4929
Score: 1821.00 Matches: 351

```

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0
US-10-698-228-1 (1-351) x PCT-US05-22501-1545 (1-4929)			
Qy	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer	20
Db	245	ATGGCGACCCGGAAGCGCGGCGGCGGATCAGGATGAGATCATCTTCA	304
Qy	21	AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg	40
Db	305	GACACCAACGAAGTGAAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTCTTCGC	364
Qy	41	ArgPheValIlePheProIleGlnIleTyrProAspIleTyrPlyMetTyrLysGlnAlaGln	60
Db	365	CGGTTTGTGTCATCTTCCAAATCCAGTACCCCTGATATTTGGAAAAATGATAAACAGGCACAG	424
Qy	61	AlaSerPheTrpThrAlaGluGluValAspLysSerLysAspLeuProHisTrpAsnLys	80
Db	425	GCITCCTTCTGGACAGCAGAAGAGTTCGACTTATCAAAAGGATCTCCCTCACTGGACACAAG	484
Qy	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaIleAspAsp	100
Db	485	CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCCCTTTTTCAGCCAGCATGTAT	544
Qy	101	GlyIleValAsnGluAsnLeuValGluAtrGpSerGlnGluValGlnValProGluAla	120
Db	545	GGAAATTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCAGAGGCT	604
Qy	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140
Db	605	CGCTGTTCTATGGCTTTCAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG	664
Qy	141	LeuIleAspThrTyrIleArgAspProLysArgGluPheLeuPheAsnAlaIleGlu	160
Db	665	CTGATGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA	724
Qy	161	ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys	180
Db	725	ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCCCTTGCAGTAGCAGATAGAAAA	784
Qy	181	SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly	200
Db	785	TCTACTTTTGGGAAAGAGTGGTGGCTTGTCTGTAGAAAGAGTTCCTCTCTCAGGA	844
Qy	201	SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer	220
Db	845	TCITTTTGTCTATATTTCTGGCTAAAGAGAGAGGCTTATATGCGAGACTCACTTTTTC	904
Qy	221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln	240
Db	905	AATGAACTCATACGACAGATGAAGACATCACTGTGACTTTCCTGCTCGCTGATGTTCCAA	964
Qy	241	TyrLeuValAsnLysProSerGluGluAtrGValArgGluIleIleValAspAlaValLys	260
Db	965	TACTTAGTAATTAAGCCTTCAGAAAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA	1024
Qy	261	IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle	280
Db	1025	ATTGAGCAGCAGTTTTAAACAGAAGCCTTGCCAGTTGGCCCTCATTTGGAATGTAATTCAT	1084
Qy	281	LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer	300
Db	1085	TTGATGAACAGTACATTTGATTTGTAGCTGCAGATTACTTGTGGAACCTTGATTTCTCA	1144
Qy	301	LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys	320
Db	1145	AAGTTTTCAGCGAGAAATTCCTTTTGATTTTATGGAACAACTTCTTTTAGAGGAAAA	1204
Qy	321	ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr	340

Db 1205 ACAAATTTCTTTGAGAAACAGCTTTTCAGAGTATCAGCGTTTTTGCAGTTATATGCGAGAAACC 1264
Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1265 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1297

RESULT 4

PCT-US05-09639-44
; Sequence 44, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-44

Alignment Scores:
Pred. No.: 2,14e-181 Length: 4955
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-10-698-228-1 (1-351) x PCT-US05-09639-44 (1-4955)

Qy 1 MetGlyAspProGluArgProGluAlaLaglyLeuAspGlnAspGluArgSerSer 20
Db 245 ATGGCGAGCCCGAAGCGCGAAGCGCGCGGCTGGATCAGATGAGAGATCATCTTCA 304
Qy 21 AspThrAsnGluSerGluLeuIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
Db 305 GACACCAACGAAGTGAATAAGTCAATGATGAAGAGCCACTCTTAAGAAAGATTTCTCGC 364
Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
Db 365 CGTTTGTTCATCTTCCATCCAGTACCTCTGATATTTGGAAATGATATAACAGGCACAG 424
Qy 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
Db 425 GCTTCTCTTCGACAGCAGAGAGTTCGACTTATCAAGGATCTCCCTCACTGGAACAAG 484
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 485 CTAAAGCAGATGAGAAGTACTTCATCTCTCATCTTAGCCCTTTTTCGACGCACTGAT 544
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 545 GGAATTTAAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 604
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 605 CGCTGTTCTATGGCTTTCAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 664
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 665 CTGATAGACACTTACATCAGAGATCCAAAGAAAGGGAATTTTATTTAATGCAATTGAA 724
Qy 161 ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
Db 725 ACCATGCCCTATGTAAAGAAAGAGAGATGGGCTTGGATGGATAGCAGATAGAAA 784
Qy 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200

Db 785 TCTACTTTTGGGAAAGAGTGGCGCTTCTGCTCTGTAAGAGGAGTTTTCTTCTCAGA 844
Qy 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 845 TCTTTTGTCTATATTTCTGGCTTAAGAAAGAGAGTCTTATGCGAGACTCCTTTTCC 904
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 964
Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 965 TACTTTAGTAAATAGCCTTCAGAAAGAAAGGCTCAGGAGATCATTTGTGTGCTCAAA 1024
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 1025 ATTGACGAGGAGTTTTTAACAGAACCTTCCAGTGTGGCTCATTTGGAATGATTCAT 1084
Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGlyPheSer 300
Db 1085 TTGATGAACAGTACATTTGATTTGTAGCTGACAGATTACTTTGGAAACTTTGGATTCTCA 1144
Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGlyLys 320
Db 1145 AAGTTTTTTCAGGCAGAAAAATCCTTTTGTATTTATGGAACAATTTCTTTAAGAGGAAA 1204
Qy 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 1205 ACAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGGTTTTTGCGATTATGCGAGAAACC 1264
Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1265 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1297

RESULT 5

PCT-US05-21650-13
; Sequence 13, Application PC/TUS0521650
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: MIGFS AS MODIFIERS OF THE IGF PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX05-017
; CURRENT APPLICATION NUMBER: PCT/US05/21650
; CURRENT FILING DATE: 2005-06-27
; PRIOR APPLICATION NUMBER: 60/581,689
; PRIOR FILING DATE: 2004-06-21
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-21650-13

Alignment Scores:
Pred. No.: 2,14e-181 Length: 4955
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-10-698-228-1 (1-351) x PCT-US05-21650-13 (1-4955)

Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 245 ATGGCGAGCCCGAAGCGCGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
Qy 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
Db 305 GACACCAACGAAGTGAATAAGTCAATGATGAAGAGCCACTCTTAAGAAAGATTTCTCGC 364
Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60

```
Db 365 CGTTTGTTCATCTTTCCCAATCCAGTACCTCGATATTTGGAAAATGTATAAACAGGCACAG 424
Qy 61 AlaSerPheThrAlaGluGluValAlaSerLysLeuSerLysAspLeuProHisTrpAsnLys 80
Db 425 GCTTCTCTGACAGCAGAGAGGTGCTTATCAAGGATCTCTCTCCTCAGTGGAAACAAG 484
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 485 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCATCTTCTTCTTCTTCTTCTTCTTCTTCT 544
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 545 GGAATTTGTAATTAAGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCCAGAGGCT 604
Qy 121 ArgCysPheThrGlyPheGlnIleLeuLysGluValHisSerGluMetTyrSerLeu 140
Db 605 CCGTGTCTTATGGCTTTCAAAATCTCATCGAGAATGTTCTCATCGAGATGTACAGTTTG 664
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 724
Qy 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCGCTTGGCATGGATGGATAGAAAA 784
Qy 181 SerThrPheGluValAlaPheAlaValAlaGluValGluValGluValPhePheSerGly 200
Db 785 TCTACTTTTGGGGAAGAAGTGGTGGCTTTGCTGCTGTAGAAGAGATTTTCTCTCAGGA 844
Qy 201 SerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 845 TCTTTTGTCTATATTTCTGGCTTAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 904
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCTCATCTGTGACTTTGCTTGGCTGATGTTCAA 964
Qy 241 TyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAlaValLys 260
Db 965 TACTTAGTAAATTAAGCCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAA 1024
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 1025 ATTGAGCAGGAGTGTATTAACAGAAAGCCTTGCAGTGGCGCTCATTTGGAATGAATGCAAT 1084
Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGlyPheSer 300
Db 1085 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGTTGGAACCTTGGATTTCA 1144
Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 1145 AAGGTTTTTTCAGGCAGAAAAATCCTTTGATTTATGAAAGAACATTTCTTTAGAGGAAAA 1204
Qy 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 1205 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTGCGAGTTATGCGAAGAAC 1264
Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1265 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1297
```

RESULT 6

```
US-11-088-686-44
; Sequence 44, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; PRIORITY FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
```

```
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-088-686-44
```

```
Alignment Scores: 2.14e-181 Length: 4955
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Dbs: 27 Gaps: 0
```

US-10-698-228-1 (1-351) x US-11-088-686-44 (1-4955)

```
Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
Db 245 ATGGGCGACCCGGAAGGCCGGAAGCGCGCGCTGGATCAGGATGAGAGATCATCTTCA 304
Qy 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
Db 305 GACACCAACGAAAGTGAATATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Db 365 CGGTTTGTTCATCTTTCCAATCCAGTACCTTGATATTTGGAAAAATGTATAAACAGCACAG 424
Qy 61 AlaSerPheThrAlaGluGluValAlaSerLysLysAspLeuProHisTrpAsnLys 80
Db 425 GCTTCTCTTGGACAGCAGAGAGTGGCTTATCAAGAGGATCTCCCTCCTCCTGGAAACAG 484
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 485 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCATCTTTAGCCTTTTTCAGGCCAGTGTAT 544
Qy 101 GlyIleValAsnGluAsnLeuValArgPheSerGlnGluValGlnValProGluAla 120
Db 545 GGAATTTGTAATTAAGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGTTCCAGAGGCT 604
Qy 121 ArgCysPheThrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 605 CCGTGTCTTATGGCTTTCAAAATTTCTCATCGAGAATGTTCTCATCGAGATGTACAGTTTG 664
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 724
Qy 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCGCTTGGCATGGATGGATAGAAAA 784
Qy 181 SerThrPheGluValAlaPheAlaValAlaGluValGluValGluValPhePheSerGly 200
Db 785 TCTACTTTTGGGGAAGAAGTGGTGGCGCTTTGCTGCTAGAAGAGATTTTCTTCTCAGGA 844
Qy 201 SerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 845 TCTTTTGTCTATATTTCTGGCTTAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 904
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCTCATCTGTGACTTTGCTTGGCTGATGTTCAA 964
Qy 241 TyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAlaValLys 260
Db 965 TACTTAGTAAATTAAGCCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAA 1024
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 1025 ATTGAGCAGGAGTGTATTAACAGAAAGCCTTGCAGTGGCGCTCATTTGGAATGAATGCAAT 1084
```

Qy	281	LeuMetLysGlnTyrlleGluPheValAlaAspArgLeuValGluLeuGlyPheSer	300
Db	1085	TTGATGAAACAGTACATTGTAGTCTGACAGATTACTTGTGGAACTTGGATTCTCA	1144
Qy	301	LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys	320
Db	1145	AAGTTTTTTCAGGCAGAAATCCCTTTTGATTTATGGAAACATCTCTTTAGAAAGGAAAA	1204
Qy	321	ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr	340
Db	1205	ACAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGCGCAAAACC	1264
Qy	341	ThrAspAsnValPheThrLeuAspAlaAspPhe	351
Db	1265	ACAGATAACGCTTTTCACCTTGGATGTCAGATTTT	1297

RESULT 7

```

PCT-US05-09639-18
; Sequence 18, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004WO1
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-18

```

Alignment Scores:	
Pred. No.:	7,97e-182
Score:	1817.00
Percent Similarity:	100.00%
Best Local Similarity:	99.72%
Query Match:	99.78%
DB:	1
Length:	1056
Matches:	350
Conservative:	1
Mismatches:	0
Indels:	0
Gaps:	0

US-10-698-228-1 (1-351) x PCT-US05-09639-18 (1-1056)

Qy	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer	20
Db	1	ATGGCGCACCCGGAAAGCCGGAGCGCGCGCTGGATCAGATGAGAGATCATCTTCA	60
Qy	21	AspThrAsnGluSerGluLeuIeIysSerAsnGluGluProLeuLeuArgIysSerAsn	40
Db	61	GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCCTTAAGAAGAAGTTCTCGC	120
Qy	41	ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln	60
Db	121	CGGTTTGTTCATCTTCCAAATCCAGTACCCCTGATATTGGAAAATGATATAACAGGCACAG	180
Qy	61	AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys	80
Db	181	GCCTTCCTCTGGACAGACAGAGAGTGCACCTTATCAAAAGGATCTCCCTCAGTCGAACAG	240
Qy	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaIaSerAsp	100
Db	241	CTTAAGCAGATGAGAAATGATTCATCTCTCACATCTTAGCCCTTTTTCAGCCCAAGTGA	300
Qy	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla	120
Db	301	GGAAATTGTAANTGAAAAATTTGGTGAGCGCTTATCAGGAGTGCAGGTTCCAGAGGCT	360
Qy	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140
Db	361	CGCTGTCTCTTTGGCTTTTCAAAATCTCATCGAGATGTTTCATCTCAGAGATGATACAGTTTG	420

Qy	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160
Db	421	CTGATACAGACTTTACATACAGATCCCAAGAAAGGGAATTTTTTATTTAATGCAATTGAA	480
Qy	161	ThrMetProTyrValLysLysAlaAspTAlaLeuArgTrpIleAlaAspAqLys	180
Db	481	ACCATGCCCTATGTTTAAGAAAAGACAGATTGGCCCTTGGCATGGATAGCAGATAGAAAA	540
Qy	181	SerThrPheGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly	200
Db	541	TCTACTTTTGGCGAAAGAGTGTGGCTTTGCTGCTGTAGAAAGAGTTTCTTCTCTCAGGA	600
Qy	201	SerPheAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer	220
Db	601	TCCTTTTGCTGCTATATTCCTGGCTAAAGAGAGAGGTCTTATGTCAGAGACTCACTTTTTC	660
Qy	221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln	240
Db	661	AATGAATCATACGACAGATGAAGACTTCACGTGTGCTTGGCTGTGATGTTCCAA	720
Qy	241	TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys	260
Db	721	TACTTAGTAATAAGGCCTTCAGAAGAAAAGGCTCAGGAGATCATGTGTGATGCTGCAAA	780
Qy	261	IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle	280
Db	781	ATTGAGCAGAGAGTTTTTAACAGAAAGCTTGCCAGTTGGCCCTCAATTGAAATGAATTGCATT	840
Qy	281	LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer	300
Db	841	TTGATGAACAGTACATTGAGTTTGTAGCTGCAGAGATTACTTCTGGAACTTGGATTCTCA	900
Qy	301	LysValPheGlnalagluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys	320
Db	901	AAGGTTTTTTCAGGCAGAAAATCCCTTTTGATTTTATGAAACCAATTTCTTTAGAAAGGAAA	960
Qy	321	ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr	340
Db	961	ACAAATTTCTTTGAGAAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC	1020
Qy	341	ThrAspAsnValPheThrLeuAspAlaAspPhe	351
Db	1021	ACAGATAACGCTTTCCACTTGGATGACAGATTTT	1053

RESULT 8

```

Result 8
US-11-088-686-18
; Sequence 18, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-088-686-18

Alignment Scores:
Pred. No.: 7,97e-182 Length: 1056
Score: 1817,00 Matches: 350
Percent Similarity: 100,00% Conservative: 1
Best Local Similarity: 99,72% Mismatches: 0
Query Match: 99,78% Indels: 0
DB: 27 Gaps: 0

US-10-698-228-1 (1-351) x US-11-088-686-18 (1-1056)

```

```
Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 1 ATGGCGAGCCCGGAAGCCGGAAGCGCGCGCTGGATCAGATGAGATCATCTTCA 60

Qy 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
Db 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120

Qy 41 ArgPheValIlePheProIleGlnTrpProAspIleTrpLysMetTyrLysGlnAlaGln 60
Db 121 CGGTTTGTTCATCTTTCCAAATCAGTACCCTGATATTGGAAATGTATTAACAGGCACAG 180

Qy 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db 181 GCTTCCTCTTGACACAGAGAGAGTGCATCTATCAAGGATCTCCCTCAGTGAACAG 240

Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCAGATCTTTTTCAGCCAGTGTAT 300

Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360

Qy 121 ArgCysPheTyrGlyPheGlnIleLeuLysValAsnValHisSerGluMetTyrSerLeu 140
Db 361 CGCTGTTTCTTTGGCTTTTCAATCTCATCGAGAATGTTCTCATGAGAGTGTACAGTTTG 420

Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 421 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480

Qy 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 481 ACCATGCCCTATGTTAAGAAAGAGAGTGGCGCTTCGATGATGATAGACATGAGAAA 540

Qy 181 SerThrPheGlyGluArgValAlaPheAlaValAlaValGluGlyValPhePheSerGly 200
Db 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTTGCTGCTGAGAGAGTATTTCTTCTCAGGA 600

Qy 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGTCTATATTTCTGCTAAGAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660

Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720

Qy 241 TyrIleValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 721 TACTTAGTAAATAGCCCTTCAGAGAAAGGGTCAAGAGATCATGTTGATGCTGTCAAA 780

Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 781 ATTGACGAGGAGTTTTHAAGAGAGCCCTTGCCAGTTGGCCTCATTTGGAATGAATTCAT 840

Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 841 TTGATGAACACAGTACATTGAGTTGTAGCTGACAGATTTACTTGTGGAACCTTGATTTCA 900

Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 901 AAGGTTTTTTCAGGCAGAGAAATCCTTTTGAATTTATGAGAAACATTTCTTTAGAGGAAA 960

Qy 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 961 ACAATTTCTTTTGAGAACAGAGTTTCAGAGATATCAGCGTTTTCGAGATTATGGCAGAAC 1020

Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1021 ACAGATAACGCTTTCACCTCTGGATGCAGATTTT 1053
```

RESULT 9

```
PCT-US05-09639-20
; Sequence 20, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-20
```

Alignment Scores:

Pred. No.:	1,02e-181	Length:	1056
Score:	1816.00	Matches:	350
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.72%	Mismatches:	0
Query Match:	99.73%	Indels:	0
DB:	1	Gaps:	0

US-10-698-228-1 (1-351) x PCT-US05-09639-20 (1-1056)

```
Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 1 ATGGCGAGCCCGGAAGCCGGAAGCGCGCGCTGGATCAGATGAGATCATCTTCA 60

Qy 21 AspThrAsnGluSerGluLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
Db 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120

Qy 41 ArgPheValIlePheProIleGlnTrpProAspIleTrpLysMetTyrLysGlnAlaGln 60
Db 121 CGGTTTGTTCATCTTTCCAAATCAGTACCCTGATATTGGAAATGTATTAACAGGCACAG 180

Qy 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db 181 GCTTCCTCTTGACACAGAGAGTGCATCTTATCAAGGATCTCCCTCAGTGAACAG 240

Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCAGATCTTTTTCAGCCAGTGTAT 300

Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360

Qy 121 ArgCysPheTyrGlyPheGlnIleLeuLysValAsnValHisSerGluMetTyrSerLeu 140
Db 361 CGCTGTTTCTTTGGGCTTTTCAATCTCATCGAGAATGTTCTCATGAGAGTGTACAGTTTG 420

Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 421 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480

Qy 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 481 ACCATGCCCTATGTTAAGAAAGAGAGTGGCGCTTCGATGATGATAGACATGAGAAA 540

Qy 181 SerThrPheGlyGluArgValAlaPheAlaValAlaValGluGlyValPhePheSerGly 200
Db 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTTGCTGCTGAGAGTATTTCTTCTCAGGA 600

Qy 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGTCTATATTTCTGCTAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
```

QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 DB 661 AATGAACATCATCAGCAGAGATGAAGACATTCACCTGTGACCTTGGCTGATGTTCCAA 720
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 DB 721 TACTTAGTAATAAGCCTTCAGAAAGAAAGGCTCAGGAGATCATCTGTGATGCTGTCAAA 780
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 DB 781 AITGAGCAGAGATTTTAAACAGAAAGCCTTGCCAGTGTGCTCATTTGAATGAATGCAAT 840
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 DB 841 TTGATGAAACAGTACATTCAGTTGTAGCTGACAGATTAATCTGTGAACTTTGGATTTCTCA 900
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 DB 901 AAGGTTTTTTCAGGCAGAAAAATCCTTTTGAATTTTATGGAACAATTTCTTTAGAAAGGAAA 960
 QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 DB 961 ACAATTTCTTTGAGAAAACGATTTTCAGAGTATCAGCGTTTTCAGATTTATGCGAGAAACC 1020
 QY 341 ThrAsnValPheThrLeuAspAlaAspPhe 351
 DB 1021 ACAGATAACGTTCTTACCTTGGATGAGATTTT 1053

RESULT 10

US-11-088-686-20
 ; Sequence 20, Application US/11088686
 ; GENERAL INFORMATION:
 ; APPLICANT: Yen, Yun
 ; TITLE OF INVENTION: DRUG SCREENING
 ; FILE REFERENCE: 14037-004001
 ; CURRENT APPLICATION NUMBER: US/11/088,686
 ; PRIOR FILING DATE: 2005-03-23
 ; PRIOR APPLICATION NUMBER: US 60/556,836
 ; PRIOR FILING DATE: 2004-03-25
 ; NUMBER OF SEQ ID NOS: 59
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 20
 ; LENGTH: 1056
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-11-088-686-20

Alignment Scores:

Pred. No.: 1 02e-181 Length: 1056
 Score: 1816.00 Matches: 350
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.72% Mismatches: 0
 Query Match: 99.73% Indels: 0
 DB: 27 Gaps: 0

US-10-698-228-1 (1-351) x US-11-088-686-20 (1-1056)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
 DB 1 ATGGGCGACCCGGAAGGCGGAAGCGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
 QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
 DB 61 GACACCAACGAAGTGAATAAGTCAATGAAGAGCCACTCTCAAGAAAGAGTTCTCGC 120
 QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetLysGlnAlaGln 60
 DB 121 CGTTTGTTCATCTTCCATCCAGTACCCTTGATTTTGGAAATGATATAACAGGCACAG 180
 QY 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 DB 181 GCTTCCTCTTGACACAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAACAAG 240

QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
 DB 241 CTTAAACAGATGAGAAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT 300
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 DB 301 GGAATTCGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTCAGGTTCCAGAGGCT 360
 QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 DB 361 CGCTGTTTCTGGGCTTTCAAATTTCTATCAGAGAAATGTTCTACTCAGAGATGATACATTTG 420
 QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
 QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 DB 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCTTGGCATGGATGATAGATAGAAAA 540
 QY 181 SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
 DB 541 TCTACTTTTGGGAAAGAGTGGCTTGTCTGTAGAGAGGATTTTCTTCTCAGGA 600
 QY 201 SerPheAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 DB 601 TCTTTTGTCTATATTTCTGGCTAAGAGAGAGGCTTATGCCAGGACTCATTCTTTCC 660
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 DB 661 AATGAACATCATCAGCAGATGAAGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 720
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 DB 721 TACTTAGTAATAAGCCTTCAGAAAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAAA 780
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 DB 781 ATTGAGCAGAGATTTTAAACAGAAAGCCTTGCAGTGTGCGCTCATTTGGAATGAATGCAAT 840
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 DB 841 TTGATGAAACAGTACATTCAGTTGTAGCTGACAGATTAATCTGTGAACTTTGGATTTCTCA 900
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 DB 901 AAGGTTTTTTCAGGCAGAAAAATCCTTTTGAATTTTATGGAACAATTTCTTTAGAAAGGAAA 960
 QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 DB 1021 ACAGATAACGTTCTTACCTTGGATGAGATTTT 1053

RESULT 11

US-10-990-328A-49
 ; Sequence 49, Application US/10990328A
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele
 ; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
 ; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
 ; TITLE OF INVENTION: USES THEREOF
 ; FILE REFERENCE: CL001495
 ; CURRENT APPLICATION NUMBER: US/10/990,328A
 ; CURRENT FILING DATE: 2004-11-17
 ; NUMBER OF SEQ ID NOS: 55824
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 49
 ; LENGTH: 1849
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens

US-10-990-328A-49

Alignment Scores:

Pred. No.: 2,648-181 Length: 1849
Score: 1815.00 Matches: 350
Percent Similarity: 99.72% Conservative: 0
Best Local Similarity: 99.72% Mismatches: 1
Query Match: 99.67% Indels: 0
DB: 15 Gaps: 0

US-10-698-228-1 (1-351) x US-10-990-328A-49 (1-1849)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 85 ATGGCGACCCGGAAGCGCGAAGCGCGCGGCTGGATCAGATGAGATCATCTTCA 144
QY 21 AspThrAsnGluSerGluLeuLeuSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 145 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 204
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
DB 205 CGGTGTGTCATCTTCCATCCAGTACCCTGATATTGGAAATGTATATAACAGGCACAG 264
QY 61 AlaSerPheThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
DB 265 GTTCTCTCTGACAGCAGAGAGGTGCTATCATAGGATCTCCTCAGTGGAAACAG 324
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
DB 325 CTAAAGCAGATGAGAGTACTTCTCATCTTACATCTTAGCCTTTTTCAGCCAGTAT 384
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 385 CGAATTTGTAATGAAATTTGGTGGAGCGCTTAGTTCAGGAGGTGCAGGTTCCAGAGGT 444
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 445 CGCTGTTCTATGGCTTTCAAAATCTCATCGAAGATGTTCTCATGAGATGTACAGTTG 504
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 505 CTGATAGACATTCATACGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 564
QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
DB 565 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTTCGATGGATAGCAGATAGAAAA 624
QY 181 SerThrPheGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
DB 625 TCTACTTTTGGGGAAGAGTGTGGCTTGTCTGTAGAAGGAGTTTCTTCTCAGGA 684
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 685 TCTTTTGTCTATATTCTGGCTAAGAGAGAGGTCTTATGCCAGGACTCATTCTTCC 744
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 745 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTCCAA 804
QY 241 TyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAlaValLys 260
DB 805 TACTTAGTAATAAGCCCTTCAGAAAGAAAGGGTCAGGAGATCATATTGTGATGCTGTCAA 864
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 865 ATTGACAGAGAGTTTAAACAGAACCTTGCAGATTGGCTCATTTGGAATGAATGCAAT 924
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 925 TTGATGAAACAGTACATTGAGTTGTAGCTGACAGATTACTTTGTGGAACCTTGGATTCTCA 984
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320

DB 985 AAGGTTTTTCAGGCAGAAAAATCCTTTTGAATTTTATGAAAAACATTTCTTTAGAGCAAAA 1044
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 1045 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGAGAAACC 1104
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1105 ACAGATAACGCTCTTCACCTTGGATGCAGATTTT 1137

RESULT 12

PCT-US05-09639-22
; Sequence 22, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004WO1
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-22

Alignment Scores:

Pred. No.: 2,116-181 Length: 1056
Score: 1813.00 Matches: 349
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.43% Mismatches: 0
Query Match: 99.56% Indels: 0
DB: 1 Gaps: 0

US-10-698-228-1 (1-351) x PCT-US05-09639-22 (1-1056)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 1 ATGGCGACCCGGAAGCGCGAAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
DB 61 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysGlnAlaGln 60
DB 121 CGGTGTGTCATCTTCCAAATCCAGTACCCTGATATTTGGAAATGTATATAACAGGCACAG 180
QY 61 AlaSerPheThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
DB 181 GCTTCTCTTGGACAGCAGAGAGGTGCTTATCAAGAGATCTCCCTCATCTGGACACAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
DB 241 CTAAAGCAGATGAGAGTACTTCTCATCTTACATCTTAGCCTTTTTCAGCCAGTATGAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAGGTTCCAGAGGT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTTTCTTGGCTTTCAAAATCTCATCGAAGATGTTCTCATGAGATGTTTAGTTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTCAG 480
QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180

Db 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCCCTTCGATGCGATAGAGAAAA 540
Qy 181 SerThrPheGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
Db 541 TCTACTTTTGGGGAAGAGTGTGGCTTTGCTGTGTAAGAGGAGTTTCTTCAGGA 600
Qy 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGTGCTATATTCTGGCTAAGAGAGAGGTCCTTATGCCAGGACTCATTTTTCC 660
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Qy 241 TyrLeuValAsnLysProSerGluArgValArgGluIleValAspAlaValLys 260
Db 721 TACTTAGTAAATTAAGCCTTCAGAAGAAAGGGTCAGGAGATCATTTGATGCTGTCAAA 780
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 781 ATTAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGATGAATGTCATT 840
Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 841 TTGATGAACACGTACATTGAGTTTGTAGCTGACAGATTACTTTGTGAACCTTGGATTTCTCA 900
Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 901 AAGGTTTTTTCAGGAGAAAAATCCTTTGATTTTATGGAAGAACATTTCTTTAGAGGAAAA 960
Qy 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 961 ACAATTTCTTTGAGAAACAGGTTTCAGAGTATCAGCGTGGATCAGAGATCATCTTCA 1020

Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

Db 1021 ACAGATAACGCTCTCACCTTGGATGAGATTTT 1053

RESULT 13

US-11-088-686-22

; Sequence 22, Application US/11088686

; GENERAL INFORMATION:

; APPLICANT: Yen, Yun

; TITLE OF INVENTION: DRUG SCREENING

; FILE REFERENCE: 14037-004001

; CURRENT APPLICATION NUMBER: US/11/088,686

; CURRENT FILING DATE: 2005-03-23

; PRIOR APPLICATION NUMBER: US 60/556,836

; PRIOR FILING DATE: 2004-03-25

; NUMBER OF SEQ ID NOS: 59

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 22

; LENGTH: 1056

; TYPE: DNA

; ORGANISM: Homo sapiens

US-11-088-686-22

Alignment Scores:

Pred. No.: 2,11e-181 Length: 1056
Score: 1813.00 Matches: 349
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.43% Mismatches: 0
Query Match: 99.56% Indels: 0
DB: 27 Gaps: 0

US-10-698-228-1 (1-351) x US-11-088-686-22 (1-1056)

Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20

Db 1 ATGGGCGACCCGGAAGGCGGAGCGGCGGCTGGATCAGGATCAGAGATCATCTTCA 60

Qy 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40

Db 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Db 121 CGGTTTGTCTACTCTTCCAATCCAGTACCCCTGATATTTGGAAAAATGTATAAAGCAGCA 180
Qy 61 AlaserPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db 181 GCTTCTCTTGGACAGCAGAGAGAGGTGCACTTATCAAGAGGATCTCCCTCACTCGAACAG 240
Qy 81 LeuLeuAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT 300
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 301 GGAATTTGTAATAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGTTCCAGAGCT 360
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 361 CGCTGTTTCTTGGCTTTCAAATTTCTCATTCGAGAAATGTTCTCATCAGAGATGTTAGTTG 420
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTGAA 480
Qy 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 481 ACCATGCCCTATGTTAAGAAAAAGCAGATGCGCCTTGCAGATGGATAGAGATAGAAAA 540
Qy 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAAGAGTTTCTTCTCAGGA 600
Qy 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGTGCTATATTCTGCTTAAAGAGAGAGGCTCTATGCCAGGACTCACATTTTTC 660
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 720
Qy 241 TyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAlaValLys 260
Db 721 TACTTAGTAAATTAAGCCTTCAGAAGAAAGGCTCAGGAGATCATTTGATGCTGTCAAA 780
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 781 ATTAGCAGGAGTTTAAACAGAGCCTTGCAGTTCGCTTGCCTGATGTTTCCAA 840
Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 841 TTGATGAACACGTACATTGAGTTTGTAGCTGACAGATTACTTTGTGGAACCTTGGATTTCA 900
Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 901 AAGGTTTTTTCAGGAGAAAAATCCTTTTGTATTTATGGAAGAACATTTCTTTAGAGGAAA 960
Qy 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 961 ACAATTTCTTTGAGAAACAGGTTTCAGAGTATCAGCGTGGATCAGAGATCATCTTCA 1020
Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1021 ACAGATAACGCTCTCACCTTGGATGAGATTTT 1053

RESULT 14

PCT-US05-09639-24

; Sequence 24, Application PC/TUS0509639

; GENERAL INFORMATION:

; APPLICANT: Yen, Yun

; TITLE OF INVENTION: DRUG SCREENING

Db	721	TACTTAGTAATAATAGCCCTTCAGAAGAAAGGGTCAGGAGATCATATTGTGATGCCTGTCAAA	780
Qy	261	IleGluGlnGluPheLeuThrGluAlaLeuProValcIlyLeuIleGlyMetAsnCysIle	280
Db	781	ATTGAGCAGAGTITTTTAA CAGAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTGCATT	840
Qy	281	LeuMetLysGlnTyrlleGluPheValAlaAaspArgLeuLeuValGluLeuGlyPheSer	300
Db	841	TTGATGAACACAGTACATTAGATTGTAGCTGACAGAGATTACTTGTGGAACTTGGATTCTCA	900
Qy	301	LysValPheGlnAlaGluAanProPheAspPheMetClukasnileSerieuGluGlyLys	320
Db	901	AAGGTTTTTTCAGGCAGAAAATCCCTTTTGATTTTATGGAAAAACAATTTCTTTTAGAAGGAAAA	960
Qy	321	ThrAsnPhePheGluLysArgValSerGluTyrgLnArgPheAlaValMetAlaGluThr	340
Db	961	ACAAATTTCTTTGNAGAACCGAGTTTCAGAGTATCAGCGTTTTCGCGTTTATGGCAGANAACC	1020
Qy	341	ThrAspAanValPheThrLeuAaspAlaAaspPhe	351
Db	1021	ACAGATAACGCTTTCACCTTGGATGCAGATTTT	1053
 RESULT 15 PCT-US05-09639-26 ; Sequence 26, Application PC/TUS0509639 ; GENERAL INFORMATION: ; APPLICANT: Yen, Yun ; TITLE OF INVENTION: DRUG SCREENING ; FILE REFERENCE: 14037-004WO1 ; CURRENT APPLICATION NUMBER: PCT/US05/09639 ; CURRENT FILING DATE: 2005-03-23 ; PRIOR APPLICATION NUMBER: US 60/556,836 ; PRIOR FILING DATE: 2004-03-25 ; NUMBER OF SEQ ID NOS: 59 ; SOFTWARE: FastSeq for Windows Version 4.0 ; SEQ ID NO 26 ; LENGTH: 1056 ; ORGANISM: Homo sapiens ; TYPE: DNA PCT-US05-09639-26			
 Alignment Scores: Pred. No.: 2,7e-181 Length: 1056 Score: 1812.00 Matches: 349 Percent Similarity: 100.00% Conservative: 2 Best Local Similarity: 99.43% Mismatches: 0 Query Match: Indels: 0 DB: Gaps: 0			
 US-10-698-228-1 (1-351) x PCT-US05-09639-26 (1-1056)			
Qy	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAaspGlnAaspGluArgSerSer	20
Db	1	ATGGGGCACCCGGAAAGGCCGGAAGCGCGCGCTCGATCAGGATCAGAGATCATCTTCA	60
Qy	21	AspThrAsnGluSerGluIleLysSerAsnGluProLeuArgLysSerSerArg	40
Db	61	GACACCAACCAAAGTGAATTAAGTCAATGAAGCCACTCTCTAAGAAAGAGTTTCTCGC	120
Qy	41	ArgPheValIlePheProIleGlnTyrProAaspIleTrpLysMetTyrllysGlnAlaGln	60
Db	121	CGGTTTGTCATCTTTCCAATCCAGTACCCCTGATATTGGAAAAATGTATAAACAGGCACAG	180
Qy	61	AlaSerPheTrpThralaGluGluValAapLeuSerLysAaspLeuProHisTrpAanLys	80
Db	181	GCCTTCCTTCGGACAGCAGAAGAGGTGCAGCTTATCAAAAGGATCTCCCTCACTGGAAACAAG	240
Qy	81	LeuLysAlaAaspGluLysTyrlPheIleSerHisIleLeuAlaPhePheAlaAalaSerAsp	100
Db	241	CTTAAGACAGATGAGAAGTACTTCACTCTCAACATTTAGCCCTTTTTCAGCCAGTGAT	300
Qy	101	GlyIleValAsnGluAanLeuValGluArgPheSerGlnGluValGlnValProGluAla	120

Db 301 GGAATGTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 361 CGCTGTTTCGGGCTTTCAATTCATCGAAGTTCACCTCAGAGATGTTAGTTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
QY 161 ThrMetProTyrValLysLysAlaAspTTPAlaLeuArgTrpIleAlaAspArgLys 180
Db 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCCCTTGCATGCGATAGCAGATAGAAAA 540
QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db 541 TCTACTTTTGGGAAGAGTGGTGGCTTTCCTGTAGAGGAGTTTTCITTCAGGA 600
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGTCTATATTCCTGGCTAAAGAGAGAGAGTCTTATGTCAGGACTCAGCTTTTTC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAlaValLys 260
Db 721 TACTTAGTAAATAAGCCTTCAGAAGAAGGGTTCAGGAGATCAATTGTGATGCTGTCAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 781 ATTGACGAGAGTTTAAACAGAGAGCTTGCAGTTGGCTCATTTGGAATGAATTGCATT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 841 TTGATGAAACAGTACATTCAGTTGTAGCTGCAGATTACTTGTGAACTTGGATTCTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 901 AAGGTTTTTCAGCAGAGAAATCCTTTTATTTATGGAACACATTTCTTTAGAAAGGAAAA 960
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 961 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1020
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1021 ACAGATAACGCTCTTCACCTTGGATGAGATTTT 1053

Search completed: October 30, 2005, 10:06:54
Job time : 1278 secs